DOCUMENT RESUME

ED 027 657 EC 003 342

Proceedings of Conference for the Evaluation of Instructional Materials (Washington, D.C., April 5-6, 1968).
George Washington Univ., Washington, D.C. Special Education Instructional Materials Center.
Pub Date [68]
Note-80p.
EDRS Price MF-$0.50 HC-$4.10
Descriptors-Consumer Economics, Cooperative Planning, County School Systems, Equipment Evaluation, *Evaluation, Evaluation Techniques, *Exceptional Child Education, Guidelines, Industry, Information Centers, Information Dissemination, Information Retrieval, Information Systems, *Instructional Materials, Material Development, Publishing Industry, Textbook Evaluation

Conference participants, who are members in the 14 federally funded Instructional Materials Centers and the Council for Exceptional Children-Educational Resources Information Center, present papers on instructional materials. George M. Olshin discusses the challenge and opportunity of evaluation, Richard A. Dershimer speaks on evaluation and decision making, and Terry Denny presents the Educational Products Information Exchange approach to evaluation. Gary Adamson describes the Educational Modulation Center concept, while Floyd G. Hudson includes sample analysis cards in his talk on evaluation. Evaluation procedures in the Montgomery County Schools (Maryland) are discussed by Richard L. Darling; industry's role in evaluation is investigated by Morris Kaplan, who focuses on the Consumers Union, and by David R. Dorsett of Creative Playthings. (RP)

ED027657

NOV 20 1968

PROCEEDINGS

CONFERENCE
FOR THE EVALUATION
OF INSTRUCTIONAL MATERIALS

Washington, D.C.
April 5–6, 1968

Special Education Instructional Materials Center
The George Washington University

EC 003 342

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

PROCEEDINGS

TABLE OF CONTENTS

Preface . . . . Margaret H. Moss

Evaluation - A Challenge and an Opportunity . . . . George M. Olshin

I. Evaluation: Issues and Models

Evaluation and Decision Making. . . . . Richard A. Dershimer

The EPIE Approach to Evaluating Instructional Materials . . Terry Denny

II. Evaluation: Procedures

The Educational Modulation Center . . . . Gary Adamson

The Evaluation of Instructional Materials . . . . Floyd Hudson

Evaluation Procedures in Montgomery County Schools. . . . . Richard L. Darling

III. Evaluation: Industry's Role

The Consumers Union Model . . . . Morris Kaplan

Industry's Role in Evaluation and Education . . . . David R. Dorsett

PREFACE

On April 5 and 6, 1968, a nationwide conference for special educators on the evaluation of instructional materials was held at the Washington Hilton Hotel here in Washington, D. C. The conference was sponsored by the Special Education Instructional Materials Center,* Department of Special Education, George Washington University. Although our primary aim was to focus on the evaluation of materials for handicapped children, we recognized that evaluation procedures involve techniques which have generality and consequently the conference concerned itself with evaluation procedures, per se.

Those invited to the conference included members of the fourteen federally funded instructional materials centers (special education) and of CEC-ERIC which together make up the national network of instructional materials centers for special education. The general purpose of the conference was to provide this network with information and guidelines which would be pertinent to the evaluation of instructional materials. The speakers who addressed the group were chosen because of competency in one of the following areas: a) "issues and/or theories relating to evaluation," b) "the role of education and industry in evaluation," c) "specific guidelines and procedures for the evaluation of materials."

In this copy of the proceedings it was necessary to omit three of the presentations. One was primarily visual, and it was felt that the substance could not be communicated in written form. Another paper was omitted by a prior agreement that the paper would not be reproduced. The third paper was omitted because of an apparent breakdown in communication regarding intent to reproduce all presentations. Due to technical difficulties, it became necessary to omit reproducing the discussions which followed each speech. Selected speeches, after further editing, will appear in a monograph to be available at cost in the near future.

I want to thank the United States Office of Education (USOE) who encouraged us to undertake the conference, and Charles Williams and Pat Smyke of the American Institute of Research (AIR), whose assistance with some of the many details is appreciated. Special thanks go to the two editors: Carol Gross and Ellen Cramer, and to Ginny Swisher who was of assistance to them. Finally, these materials would not now be in your hands except for the diligence and typing skills of Veejay Naegle.

Sincerely,

Margaret Moss

Margaret H. Moss
Conference Chairman

---

*As of September, 1968, the George Washington Special Education Instructional Materials Center will be called the Mid-Atlantic Region Special Education Instructional Materials Center.

Evaluation - A Challenge and an Opportunity

by George M. Olshin, Ed.D.
Chief, Research Laboratories and Demonstration Branch,
Division of Research
Bureau for the Education of the Handicapped (BEH)
Washington, D. C.

It gives me great pleasure to welcome you to this conference on behalf of the sponsoring agency - the Bureau for the Education of the Handicapped of the United States Office of Education (USOE). The bureau is most appreciative to Mrs. Moss, Miss Gross and other staff members of the George Washington University Special Education Instructional Materials Center (SEIMC) for organizing this conference for the network of Instructional Materials Centers.

Since some persons in the audience have just become aware of the SEIMC program, perhaps it would be helpful to review briefly the purpose and development of the program.

The SEIMC Program

The concept of the SEIMC came into focus in 1962 when a Presidential Task Force recommended the development under the leadership of the USOE, of Special Education Instructional Materials Centers. The purpose of such IMC's was to provide special educators and other related professional personnel with ready access to valid instructional materials and information for the education of handicapped children and youth. It was recognized that a workable system should be established to assure that these materials are known, understood, and available to all teachers and supervisors of special education programs.

USOE followed the recommendation of the Task Force, and the program is now administered by the Bureau for the Education of the Handicapped.

To accomplish the objectives of this new effort a three phase program was planned.

Phase I - Service

This phase includes the acquisition of commercial and teacher-prepared instructional materials; describing, classifying, and organizing these materials; and the dissemination of materials and information to educators.

Phase II - Development

This next stage includes the evaluation of instructional materials, and the development and production of new materials on a pilot basis for experimental trial or demonstration to establish their effectiveness.

Phase III - Stimulation of Materials Production

The last step includes contacting the organizations which have the production capacity (such as sheltered workshops or commercial publishers) and encouraging them to produce materials which have been found to be effective in the research phase; and continuing consultation with producers to assure that ideas which they believe have merit are given consideration.

If the IMC's are to succeed, they must develop a close relationship with industry; this is a challenge and an opportunity. Very little is known about the actual educational value of many instructional materials. One of the functions of the centers will be to provide information, materials, and consultation to teachers, thus strengthening their ability to discriminate between poor and effective materials for use with handicapped children.

Conceptualizing and funding a program such as the SEIMC Network does not always guarantee success, and one knows that implementation is the more difficult task. I believe that several ingredients account for the success the IMC's have had to date. The first and most important is the calibre of the professional staff and other personnel who are associated with the administration and operation of the IMC's. The second is the autonomy given the centers to develop unique characteristics. The third, the ability of individual centers to appreciate the need for network coordination and to cooperate in the development of a functioning network. Another important factor is the ability to phase in activities to accomplish their objectives. For example, all centers knew that it was of prime importance to develop their service phase prior to initiating a research and development phase.

The objectives for the service phase are currently being accomplished by most IMC's. Centers have developed their physical facilities and extended their operations into the field. They are now working closely with state education agencies, local school systems, colleges and universities and residential institutions to develop associate or satellite centers.

Our dissemination activities are increasing and becoming more efficient and selective through center newsletters, acquisition lists, institutes and workshops, and through convention program activities. Our relationship with the Educational Resource Information Center - Council for Exceptional Children is a great strength.

Other programs--supported and not supported through our office--are now being coordinated with Network activities. For example, a USOE project at the University of Iowa which is developing and disseminating materials and information on instructional methods to special class teachers in Iowa, is now cooperating with the SEIMC Network.

It is significant too, that funds have been made available through a government research and demonstration program and implemented an educational activity that is giving direct help to the special education teacher in the field. BEH believes that this help will make a difference in the quality of his teaching. We hope that in the near future, he will become well informed about valid instructional materials--the tools of his trade.

Evaluation Activity

I see the activity of evaluating instructional materials as a great challenge to the IMC's. A challenge because the evaluation process is difficult and will require a great amount of the IMC resources.

Recognizing that problems would be encountered in the evaluation process, an IMC Committee was organized and initial efforts were made to coordinate the evaluation activities of the network. The committee became aware of outside resources currently involved in the process and sought to cooperate in these efforts.

The first major problem of evaluation is one of semantics. If the special education community and the publishers and instructional materials manufacturers continue to use different terms to describe the properties of instructional materials--progress will be slow. I hope that this meeting can help us to become aware of the problems of language in the instructional materials field.

Some of the questions the IMC's want answered which relate directly to the evaluation process are:

1. How to determine priorities?
2. Do we need a master plan?
3. What do we evaluate?
4. How do we evaluate?
5. Who will evaluate what?
6. How do the Centers cooperate in this task?
7. How can IMC's cooperate with publishers, and materials producers?
8. What are the cost factors in terms of money and professional energy?
9. Can we anticipate problems?
10. And finally, what should the role of USOE be?

I have described the genesis of the IMC program and the plan for the future. I know the IMC's will work hard to develop and implement this activity as they have the service phase. Our new friends are invited to help us and to share in the benefits we derive by improving special education and in turn serving as a model for general education.

Evaluation: Issues and Models

EVALUATION AND DECISION MAKING*

by Richard A. Dershimer, Ph.D.
Executive Officer, American Educational Research Association (AERA)
Washington, D. C.

This paper is to encourage and improve the sophistication of evaluations of all kinds. I am focusing on educational programs because of my background. Far too many supervisors and administrators are frustrated because they feel increasing pressure to assess their programs--or segments of them--only to be reminded by the specialists that the methodologies of evaluation are growing increasingly complex. Federal directives specify that programs must assess what benefits federal funds have produced, but research consultants explain that test scores and opinion questionnaires will not provide the answers.

The dilemma is heightened by statements like the one made by Louise Tyler (1968) that evaluation is now "imperative because curriculum and instructional materials development has become centralized." Yet, we are told that even the national curricular projects, with great educational leaders and hundreds of thousands of dollars from foundation and federal funds, have not been adequately assessed. Today's school administrators hear school boards use terms like cost effectiveness, objectives, and assessment with increasing frequency. They hear men like Scriven (1967) state, "Business firms can't keep executives or factories when they know they are not doing good work, and a society should not have to retain textbooks, courses, teachers, and superintendents that do a poor job when a good performance is possible."

The solution far too often is to "go through the motions," that is, gather some test score data, tabulate questionnaire results, obtain some testimonies, and write a flashy report; the data are not respected and only slightly used--if at all. A frequently used alternative is to postpone any serious evaluation until a fully qualified man can be found. The problem with both of these alternatives is that the potential for evaluation that exists in so many districts is overlooked, and valuable data are lost or ignored.

This paper will provide a rationale for the "let's get started" approach. It is meant to be a supporting document for the supervisor who is arguing for even more evaluation with his immediate supervisor, whether he is an assistant superintendent, superintendent, or chairman of the board. It will attempt to summarize and interpret the major disputes about evaluation that are most relevant to the local scene and to provide a plan that should help districts improve the way they now evaluate.

Definitions

Probably the most frequently used definition of evaluation, and the one which we shall accept for the purposes of this paper, is provided by Cronbach (1963).

---

*I acknowledge the generous contributions of Robert Stake, Robert Panos, Malcolm Provus, and Glenn Boerrigter, all of whom made valuable suggestions, most of which have been incorporated in the final draft.

He sees evaluation, broadly conceived, as "the collection and use of information to make decisions about an educational program." It is interesting that he uses the term information, and not the more precise term, data.

As valuable as this definition is, it begs a very critical issue. Does evaluation merely attempt to describe the situation under study, or does it attempt to judge how adequate, effective, or valuable something is? Most evaluators in the past have only described matters like how well or how poorly children achieved compared to certain norms, or how adequate the environment of the learner or the preparation of the professional was. Stake (1967) takes a firm stand on this issue and claims that "both [italics his] description and judgment are essential--in fact, they are the two basic acts of evaluation. Any individual evaluator may attempt to refrain from judging or from collecting the judgments of others. Any individual evaluator may seek only to bring to light the worth of the program. But their evaluations are incomplete. To be fully understood, the educational program must be fully described and fully judged."

Again, for the purpose of this paper, I concur with this extension of the definition by Cronbach, with one modification: the judgments of the evaluator must be limited to the question under study and kept within the boundaries for which there are supporting data. The evaluator is a technician and should have special insights and perceptions. But in nontechnical matters his judgment should be only one of many sought by the ultimate decision maker--supervisor, superintendent, or school board. He can be asked how effective a new elementary school science program is, for example. His reactions should be weighed against those of the business manager who may have been analyzing the costs of the new program, the assistant superintendent who knows the feelings of the parents about the existing elementary science program, and the principal who knows the qualifications and reaction of the teachers.

The term educational program used in this sense refers to those aspects of any school situation involving the direct interaction between a group of professionals and a group of students in a school. Educational materials can be included in this definition; budgets cannot. Organization of teaching teams can be included; administrative reorganization of the central office cannot. Evaluation of the progress of a single student or even several students by a single teacher is not included in our definition.

Evaluation is an activity or set of activities initiated or utilized to provide data for major operational decisions in the schools. It is only one of several sources of data for those major decisions even though, on matters relating to the educational program, it may be the most relevant and most significant. Evaluation serves three other purposes, however: (1) providing more systematic ways of gathering data for many other purposes; (2) sharpening the objectives of any organization; and (3) identifying and clarifying aspects of the situation where revision or improvement is most needed and/or desired.

A Schema for Evaluation

All too frequently, in the mind of the administrator, evaluation is equated with checklists, tests, questionnaires--in other words, with techniques rather than an overall plan. I am proposing a three-part plan, or schema, the three major components of which are (1) antecedent conditions or input variables, (2) intervening events or treatments, and (3) outcomes or objectives.

Ideally, no evaluation should be considered complete, and no decision made, until meaningful data from all three sources are available. But, typically, data are uneven or nonexistent. The socio-economic position of parents or their child-rearing attitudes, critical personality characteristics of the children or teachers--these are only some of the many input variables that are difficult at best to obtain. What does the pretty, young teacher down the hall do that causes her pupils to perform better than children from other classes in the same building year after year? And how can we know whether we really changed the behavior of children without following them for the next 25 years? These questions are merely illustrative of the limitation in designating intervening variables and outcome data.

Even though there may be large gaps in the data, the schema is still useful, in that it will help the evaluator and administrator understand the limitations of the data that are on hand. For example, the fact that there are no records for one-third of the pupils in a given class because their parents are migrant workers (a highly relevant piece of datum) should prevent educators in that school from generalizing whatever success they have had with a particular project.

The schema has two other advantages: (1) It should allow evaluators and educators to break out from a myriad of activities smaller bits of data on behavior (performance) or relationships between people and materials, etc., so that they will be better able to focus on those variables which are most likely to be relevant to the program under study, and (2) It should lead to the collection of more and more varied data.

The simplicity of the three-part schema is deceptive. Any one part really must be thought of as a link in an interlocking chain. Since the schema must be kept dynamic, that is, kept as part of a constantly shifting scene, the final outcomes from one schema may serve as antecedent conditions for another--or intermediate objectives for yet a third. How many links in the chain are used in a given situation obviously will depend on the degree of competence of the evaluators and the complexity of activities being studied.

Before leaving the schema, I must point out that it is a fairly common approach and first appeared in print with the New York State Quality Education Project in the middle 1950's. Many writers, of whom Stufflebeam* and McQuire (1967) are excellent representatives, advocate a fourth component--the setting or context. These teams include the identification of the problem and the assumptives that undergird the problem. While I agree with the importance of these factors, I still see them as input variables, as I point out later on. But let's examine each component of the schema in greater detail.

*See the CIPP Evaluation Model by Daniel I. Stufflebeam presented on page 35, "Columbus Report," BIG CITY TITLE I EVALUATION CONFERENCE, 1967 REPORT, Pittsburgh Public Schools.

Input Variables or Antecedent Conditions

The importance of input variables has been excellently summarized by Stake (1967):

> What one finds when he examines formal evaluation activities in education today is too little effort to spell out antecedent conditions and classroom transactions (a few of which visitation teams do record) and too little effort to couple them with the various outcomes (a few of which are portrayed by conventional test scores.)

Inputs may be defined as all the relevant characteristics that the principals involved in any educational enterprise under study bring with them at the beginning of the enterprise. Training, age, experience, level of competence are typical of the antecedent data usually gathered about teachers. But a wise evaluator will attempt to gather data about the interrelationships of teachers--the pecking order for example--, and the attitudes of teachers toward "outsiders", even if they come from the downtown, central office.

A frequently overlooked input datum is the information that led to the initial identification of the problem; evaluations have a way of becoming stereotyped. A school board member's child is having difficulty. The result may be the assessment of the performance of a single teacher rather than examination of the performance of all the children with characteristics similar to the single example. Conversely, an irate PTA meeting might lead to evaluation of the public relations program in the district rather than analysis of the educational problems highlighted by the parents. The assumptions that lead to quick interpretations must be questioned early and often.

Treatment of Intervening Variables

The data that are to be gathered in this category are from the relevant transactions that have taken place among the students, professionals, media, and materials in the situation under study. The term relevant in this case should be considered quite loosely and should depend in large measure on the type of data an evaluator wants to collect.

The role the evaluator plays in specifying the treatments frequently causes problems. Brickell (1961) points out that evaluators require two restrictions: "(1) procedures must not be changed in midstream, otherwise it will not be clear what is being evaluated; and (2) the circumstances in which the procedures are used must be kept comparable, otherwise it will not be clear what is determining the outcome." In order to move every student toward the desired outcomes, the teacher will often use any promising means, regardless of treatments specified. If a student is not able to comprehend the concept of integers, for example, the teacher will want to find a new text, or a programmed text, or extra help from another student for him.

This conflict has been and remains one of the major conflicts that prevent many schools from conducting more evaluations. I place most of the blame on the evaluators themselves, for reasons I shall discuss in more detail later on. Evaluators should become more ethnographic. That is, they should be more willing to back off and observe what treatment teachers actually employ under given circumstances with given children and attempt to formulate hypotheses or gather other data based on these observations. This somewhat heretical recommendation is made because of the continued inadequate state of knowledge about teaching and learning in formal school settings, and with the realization that most evaluations in local schools are not aimed at--at least should not be aimed at--increasing the world's knowledge of major educational issues. The concern with control has not provided the resulting trade-off in knowledge. As Gage (1963) states the case, "research in teaching, employing the most commonly accepted paradigm, has been abundant--hundreds of studies yielding thousands of correlation coefficients have been made. In the main, these studies have yielded disappointing results: correlations that are nonsignificant, inconsistent from one study to the next, and usually lacking in psychological and educational meaning."

We must be aware of the difficult position in which we place the local director of research if we do not permit him to tightly structure his designs and controls. Usually he has been prepared for his work by professors of research who continue to serve as gateposts for career opportunities. The reference groups and the professional associations to which he belongs expect him to contribute scientific papers, and they evaluate his work by the standards of research. For these reasons, it is obvious to me that persons who use research skills primarily to improve what is done within a given school system or institution badly need a separate, more clearly identifiable professional group. It should be a group whose members will be just as concerned as their scholarly counterparts with quality and quantification, but who will recognize the difference between gathering data for decisions, which is basically an engineering function, and attempting to add to knowledge, a research function.

There is yet another reason, however, why the teacher's concern for the individual frequently conflicts with the evaluator's concern for structure: individualization of instruction has not been adequately conceptualized in a way that would permit any kind of a systematic evaluation. A new approach to individualization, "mastery learning," may supply this conceptual structure. Bloom (1968) provides a clear description of the term, "teaching for mastery." He cites evidence from pilot studies that by individualizing instruction and varying the time allotted for learning tasks, and to individuals, up to 95 percent mastery becomes a goal, it follows logically that evaluators must be expected to evaluate different kinds of treatments than they have in the past. The evaluation problems of mastery learning have not yet been explored in any depth, but the approach, I posit, still holds great promise. The approach should have value both for home-grown courses and curricular innovations and for courses and innovations packaged by outside sources.

Objectives

The importance of objectives to evaluation has not been challenged since the eight-year study (see Smith and Tyler, 1942). In recent years, objectives have assumed ever greater importance as can be seen in this quote by Gagné (1965):

> For the person who wishes to study the process of education, to analyze it, to perform research upon it for the purpose of understanding and improving it, statements of educational objectives as human performance are an *absolutely essential starting point*. [italics mine].

There is an emerging dissent to the emphasis placed on objectives--particularly the need for behavioral terms. Robert Travers, one of the most persistent critics, bases his objections on two points: (1) There are too many "behaviors" like creativity and abstract reasoning that cannot be specified and measured, and these then are usually underemphasized in educational programs. (2) The behavioral objectives approach rests on the false assumption that children are like "plastic masses" (1968) of raw materials that are inexorably molded into the shape foreseen by the planner. The only conclusion (which will be of little help to the practitioner, I fear) that seems compatible with the points of view represented by Gagné on one hand and Travers on the other is that behavioral objectives should be used for those teaching situations where detailed objectives are possible, remembering that there are "higher" objectives that cannot be treated in the same way.

Another issue that must be highlighted is how to balance locally derived objectives with those established for larger populations. In a private school that sends large numbers of graduates to leading universities, the two may be synonomous. In a slum school it may be unreasonable to expect children to be judged by national norms.

This leads to the separate issue of how to evaluate the objectives that teachers adopt for themselves, or for a school. If a junior high school staff wants all students to develop a written proficiency in Swahili, proficiency defined at a fairly advanced level, the central staff should be very interested to know why. Given this kind of priority, the previous objectives should be reexamined thoroughly to see if the rationale for all other courses can still be justified.

Although I realize I have treated the concept of objectives very superficially, the literature is now so extensive that I prefer to concentrate on aspects of the schema.

The Problem of Sophistication of Data

The degree to which any evaluation will help improve decision making will depend to a large extent on how valuable the data are. Essentially there are five sources of data which may vary in the degree of sophistication and the degree to which they are empirically derived. They are: folklore, anecdotes, expert opinion, descriptive data, and research data. Each of these can help an evaluator make certain judgments and administrators make certain decisions, but each has its limitations. For this reason it is important that we examine each source of data.

Folklore

Astin and Panos (1968) call folklore "any widely accepted but empirically untested assumption concerning a causal relationship between an educational program or operation and an educational outcome." We can name many bits of folklore, but a few will suffice. The educational justification for the junior high school is largely composed of folklore about the unique requirements of early adolescents. The belief that children should not be introduced to reading until they are six is folklore (that some teachers surprisingly still believe). Some coaches also swear that intercollegiate athletics produce better citizens.

But folklores are helpful to the decision maker because they are legitimizing beliefs. Folklores enjoy wide consensus. Therefore, the professionals are able to get on the job because they have standards for action.

The single most bothersome problem about folklore in education is that it is so often substituted for empirical evidence. It is easy to see our mistakes in the past, like the reading readiness concept. Bussing students from the slums to more favored schools is a plan born from folklore and rests on untested assumptions. So long as educators and evaluators remember this fact and are willing to question these assumptions when they can, then folklore is a valuable source of information.

Anecdotes

Anecdotes, especially dramatic ones, can have profound influence on decisions. The student who stabs a teacher and the demonstration for a project that is about to be discontinued for lack of funds, are both examples of events that focus on issues and suggest courses of action. They help call to the attention of extremely busy administrators and school boards events that may not otherwise rise to command their attention.

Since the limitations of anecdotal information should be well known, the only other comment here is to enter a plea for more effective record keeping. Just as the recording of anecdotes for individual pupils can help sketch a profile of his behavior, similar records for an entire school can reveal problems where none are believed to exist. Daily or weekly logs or diaries can serve this purpose.

Expert Opinion

The problems of "outside" consultants are well documented and need not be reiterated here. Similar problems are arising with the increased prestige accorded to research specialists on the staffs of all large school systems. Knowledge about research design and methods and statistical treatment and analysis is extremely valuable but is only a means to a larger end. The opinions of these individuals must be weighed against the opinions of many others on the staffs.

Descriptive Data

Descriptive data document what is happening in a school system and provide a systematic inventory of conditions, the incidence, distribution, and, to a certain extent, the relationships of phenomena. (For a more detailed treatment see Van Dalen, 1963, Chapter 10). I am referring to regular collection of data such as attendance records, achievement test scores, broken windows, police arrests, and to ad hoc studies such as parental opinions about pending issues, adequacy of school buildings in light of possible population shifts, and percentage of students who have smoked marijuana!

One of the problems associated with descriptive data is deciding where to focus. Schools are admonished to undertake "continuous gathering of data" (AASA, 1959). State departments of education and federal agencies seemingly have insatiable appetites for information. It will do no good to rail against the trend; it is far better to automate so as to stay on top of critical demands. Schools must adopt procedures followed today by many large industries that have their record keeping so automated that only exceptions or deviations from regular reporting are fed to the decision makers. The computers handle everything else.

As valuable as descriptive data are for evaluators and administrators, they frequently are misused. Elaborate statistical treatments, multiple regression analyses, for example, seem to imply causality. (See, for example, R. C. Nichols' review of the Coleman study, Science, 154, 1966.) If the records of 10,000 students who had male teachers show just the reverse, the implications for action are clearly suggested. But the cause of the reading problems for either girls or boys still has not been established.

Only one other comment is needed about descriptive data. The results of standardized testing frequently are equated with descriptive data. Much more imagination is needed to provide insights into the inner workings of schools--at the staff and student levels. Much information is available from secondary sources and unobtrusive data sources, two terms that will be discussed later. The number and kinds of books placed on reserve shelves in the library, kinds of equipment and supplies ordered by teachers, the frequency and variety of audio-visual equipment used, are records available to any principal and should help to supplement other pieces of information about the characteristics of teaching taking place in his school.

Research Data

The term research used in the context of this paper refers to a body of procedures and methodologies borrowed from the social and behavioral sciences that allow the evaluator to gather certain kinds of empirical data that cannot be gathered in any other way. However, the focus is still on the kinds of data that will aid in decision making, and not data that will add to new knowledge. As such it concentrates on what Kerlinger (1965) calls the "shorter range goal of finding specific relations." Consequently, the researcher is restricted in the way he can generalize the results.

Astin and Panos (1968) use the criterion of replication to distinguish research procedures from descriptive procedures even though they acknowledge that it is "rarely feasible to replicate the (classical) experimental conditions on any substantial scale." To replication I prefer to add the concept of intervention, that is designing a set of experiences and testing and observing the reactions of the individuals to those experiences.

There is little agreement on the most appropriate procedures to obtain the desired data. In classroom settings it should help administrators to use comparative data whenever possible. That is, the data gathered about the behavior of students or staff should be compared both to a set of absolute standards as well as to data gathered from comparable groups. Research also is the best way to obtain clues about differentiated effects of treatments. Administrators must be concerned with discovering the various reactions of different kinds of students under differing conditions, not just with gross data like group means and correlation between means.

The problems related to the use of experimental and quasi-experimental methods are too complex to take up in this paper. (For one of the most comprehensive overviews see Campbell and Stanley, 1966.) To obtain the most sophisticated data possible, obviously a specialist in research design and methodology must be employed. School systems which do not have these specialists should either hire them or take the responsibility for retraining some of their competent young men and women. Hiring part-time consultants is a very poor substitute.

Using More Imaginative Procedures

School systems should insist that their evaluators develop and use more imaginative procedures for gathering data for decisions and not try to meet criteria from the research community necessary to add to new knowledge. There are five "rules" that should lead to richer data:

1. Use more than a single measure. One quote from Webb and others (1966) will support this point.

> So long as we maintain, as social scientists, an approach to comparisons that considers compensating error and converging corroboration from individuality contaminated outcropping, there is no cause for concern. It is only when we naively place faith in a single measure that the massive problems of social research vitiate the validity of our comparisons.

2. Use more than a single mode of measurement. The beginning researcher almost inevitably thinks in terms of pencil-and-paper tests and questionnaires. He forgets that it is possible to observe behavior, to use logs and diaries, interviews, projective techniques--to mention only a few possibilities.

3. Differ treatments among treatment groups. Again, if the emphasis is on generating meaningful data for decisions, more than one group should be used, and the treatments should be differentiated. Cronbach's appeal (1963) for the use of double blind experiments in educational research has gone largely unheeded. In other words, there should be at least one experimental group, one control group, and a third group to whom is given some kind of placebo, that is, an alternative treatment that has different ingredients but affords the third group as much attention as the other groups.

4. Make your instruments and methods sensitive to associated data and possible serendipitous interpretations. If you want to evaluate changes in behavior among teachers who are sent off to NSF summer institutes, it is important to find out not just if they teach the new materials they were taught, but what materials they no longer use. It is as important to determine the new sources of information they use and, perhaps the individuals with whom they now consult.

5. Use more secondary data, or unobtrusive measures. I have borrowed the term unobtrusive measures from a book of the same name by Webb, Campbell, Schwartz, and Sechrest (1966). They argue that if you are to use multiple measures, you just find ways of discerning behavior without artificially altering their behavior. For recording behavior they advocate using physical evidence (erosion of tile around certain exhibits in museums and archives), observation (conversation sampling and type of dress), and hidden hardware.

Using these measures opens up many exciting possibilities for evaluators in schools. Consider just one example: A principal has urged his teachers to undertake more joint planning as a way of moving his school toward team teaching. He makes it possible for them to meet during school time, but he would like to know who is really undertaking some joint planning. He can ask them, or give them a questionnaire, but he does not want to interfere, just obtain a clue as to progress.

He assumes that if teachers are going to plan future activities, they will need certain books in advance from the library (cards can be checked). They should request certain pupil records that are kept in the office in order to compare notes on individual students. As the plans unfold, they should order different materials than they have in the past and request permission to use the small auditorium for large-group meetings.

The quantification of many of these variables will be very difficult, if not impossible. However, this should not excuse any evaluator from measuring, as precisely as possible, as much data as possible. He must determine the reliability of his observations and the validity of the attempted improvements. So regardless of the level of statistical sophistication used in the beginning, the tendency should be over any period of time to use ever more sophisticated and complex measurements.

Concluding Statements

The "let's get at it approach" can be criticized in several ways. It is quite likely to generate and perhaps condone poorly conceived and sloppily executed evaluations as well as provide only evidence of the kind we have used in the past. It is an oversimplified approach to highly complex problems and may tend to focus the educators on superficial aspects of those problems.

There are two ways to compensate for these weaknesses. First develop a broad data base, focus it toward major program decisions and keep it current. Secondly, theories should be used as frameworks and as guides for the collection and interpretation of data whenever possible. Let me elaborate somewhat on the second point.

In spite of the unquestioned value of theory to the sciences, it remains alien to practitioners in education. And with some justification, since theories in our field remain weak (Bloom, 1966). Nevertheless, without theories, investigations will be random and diffused. As Kerlinger (1965) defines theory it is "a set of interrelated constructs (concepts), definitions, and propositions that presents a systematic view of phenomena by specifying relations among variables, with the purpose of explaining and predicting the phenomena."

The importance of theory was reaffirmed recently in the study of the utility of research completed by the Defense Department. In this highly mission-oriented agency it is only natural that they favor research that will lead as directly as possible to technological developments. Project Hindsight, the title of the Department of Defense study, confirmed that "focused research" does have great payoff. But it also pointed out the debt owed to theory. Sherwin and Isenson point up this fact in their summary of the study in Science (1967):

> None of our science Events (their term for discoveries) could have occurred without the use of one or more of the great systematic theories--classical mechanics, thermodynamics, electricity and magnetism, relativity and quantum mechanics. These theories also played an important role in many of the technology Events. If, for example, we were to count the number of times that Newton's laws, Maxwell's equations, or Ohm's law were used in the systems we studied, the frequencies of occurrence would be so high that they would completely overshadow any of the recent Events we identified.

Educators who doubt the value of theories should reread Robert Merton's Social Theory and Social Structure (1957). Even theories at the lowest level of generalizability give greater meaning to data and provide insights for improving the studies that follow the present investigation.

But the final argument in support of the "let's get started" approach has nothing at all to do with the technical nature of evaluation or research. It is rooted in the reality that decisions must be made in schools about program matters. They can be based on very limited data and large amounts of experience-based opinions, or just the reverse. The relationship of the evaluator to the decision maker in schools is precisely that described by Etzioni (1968):

> In short, the relationship between the social sciences and a societal decision-maker is not very different from that between the natural sciences and a medical practitioner: Even if either practitioner had mastered all knowledge which the scientific discipline contains, he still would have to interpret, project, and make connections--on the basis of fragmented information and in accordance with the canons of the applied world.

In short, the evaluator should concentrate on providing the most valid, the most reliable and the most relevant information and the best judgments concerning the information he gathers. After that it becomes the administrator's responsibility to act or not to act as he sees fit. The recognition of this differentiated responsibility should be the final, clinching argument for getting started with more formal evaluations of educational programs.

REFERENCES

American Association of School Administrators. Judging schools with wisdom. 1959.

Astin, A.W. & Panos, R.J. The evaluation of educational programs. Manuscript in process for the revision to Educational Measurements, American Council on Education, Washington, D.C., 1968.

Bloom, B.S. Mastery learning for all. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1968.

Bloom, B.S. Twenty-five years of educational research. American Educational Research Journal, 1966, 3, 211-221.

Brickell, H.M. Organizing New York State for educational change. State Education Department, Albany, New York, 1961.

Campbell, D.T. & Stanley, J.C. Experimental and quasi-experimental designs for research on teaching, New York: Rand McNally. 1966.

Cronbach, L.J. Course Improvement through evaluation. Teachers College Record, 1963, 64, 672-683.

Etzioni, A. How may congress learn? Science, 1968, 159, 170-172.

Gage, N.L. Paradigms for research on teaching. 1968, In N.L. Gage (Ed.) Handbook of Research on Teaching. New York: Rand McNally, 1963.

Gagné, R.M. Educational objectives and human performance. In Learning and the Educational Process. J.D. Krumboltz, (Ed.), New York: Rand McNally, 1965.

Kerlinger, F.N. Foundations of behavioral research. New York: Holt, Rinehart, & Winston, 1965.

McGuire, C.A. A proposed model for the evaluation of teaching. The Evaluation of Teaching, a report of the Second Pi Lambda Theta Catena, Pi Lambda Theta, Washington, D.C., 1967.

Merton, R.L. Social theory and social structure. New York: The Free Press, 1967.

Scriven, M. The methodology of Evaluation. Perspectives of Curriculum Evaluation, American Educational Research Association, 1967.

Smith, E.R., & Tyler, R.W. Appraising and recording student progress. New York: Harper & Row, 1942.

Stake, R.E. The countenance of educational evaluation. Teachers College Record, 1967. 68, 523-540.

Travers, R.W. Evaluator of new instructional procedures. Paper presented at the meeting of the American Educational Research Association, Chicago, 1968.

Tyler, L. & Klein, M.F. Recommendations for curriculum and instructional materials. Report, University of California, Los Angeles, 1967.

Van Dalen, D.B. *Understanding educational research*. New York: McGraw-Hill, 1962.

Webb, E.T., Campbell, D.T., Schwartz, R.D. & Sechrest, L. *Unobtrusive measures: Nonreactive research in the social sciences*. New York: Rand McNally, 1966.

THE EPIE APPROACH TO EVALUATING INSTRUCTIONAL MATERIALS

TERRY DENNY, Ed.D.
Head, Educational Products Information Exchange (EPIE)
Research Office, Urbana, Illinois

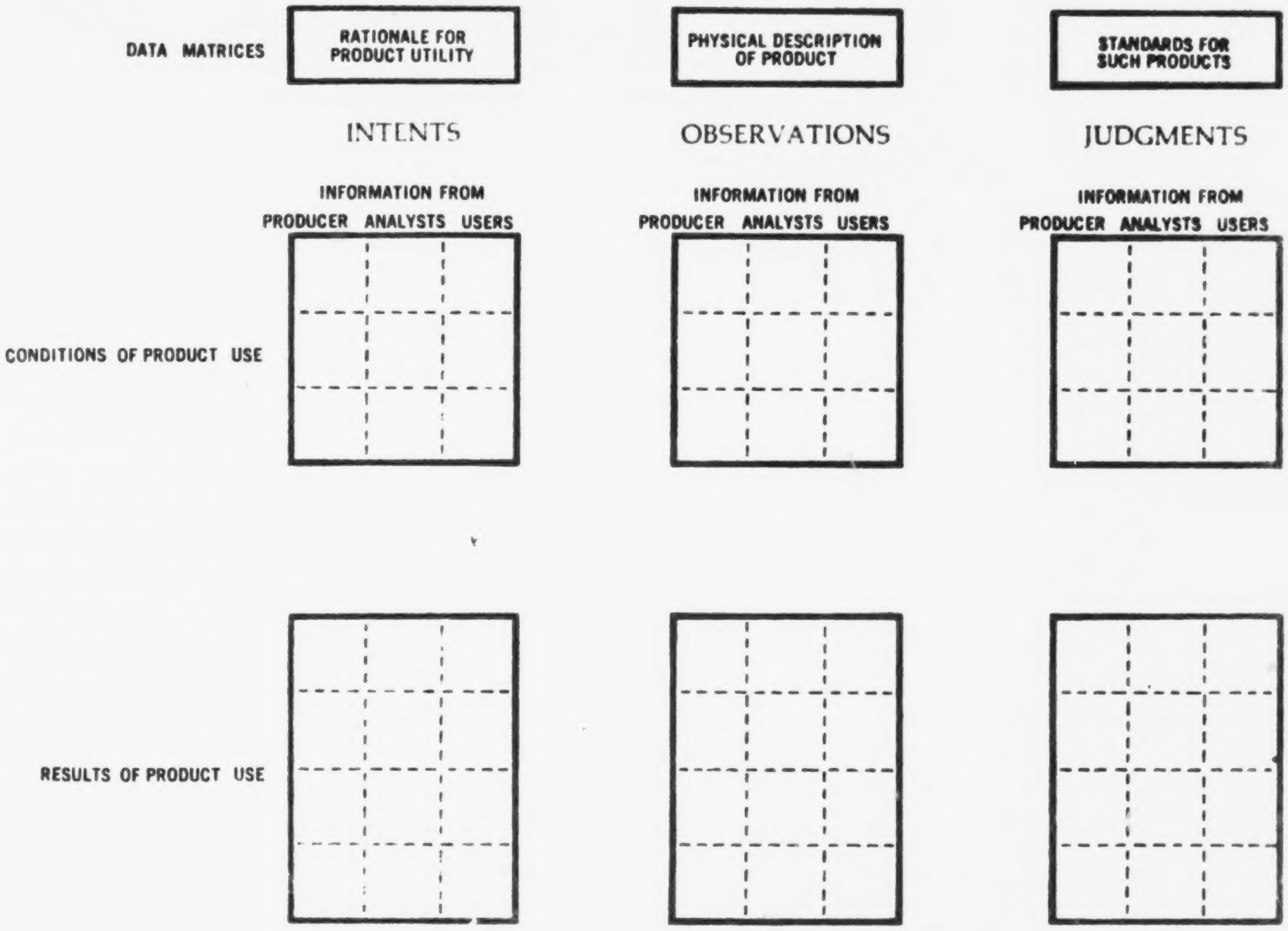
In my dark ages Robert Stake (1967a) began talking about an approach to evaluation in education, an approach designed to reveal the countenance of education. I heard about it and was interested, read about it and was intrigued, thought about it and was possessed. During my pleasant years on Purdue University's educational psychology staff I had been involved in a variety of short and long-term experiments in the name of evaluation. When the going was rough, I hugged my Linus blanket, the venerable Tyler rationale. Then, too, I had a few AERA-APA model-building papers into a special file marked "do something with these, someday" and had been dimly aware of a few elements of a way of thinking about educational evaluation that made sense to me. It was pretty chaotic. I used to answer "research" instead of "evaluation" whenever a colleague asked me what I was doing . . . no matter what I was doing. Stake's countenance paper clouted me.

About a year later I began hearing about a systematic attempt to institutionalize the evaluation of the thing-side of education: educational products. P. Kenneth Komoski (1967) was starting the EPIE Institute and had articulated a captivating set of technological propositions about what needed to be done and how it might be done. The Institute has evolved an evaluation model principally through the thoughts of Stake and Komoski, and it serves us well. Indeed, I have talked and written so extensively about it I have come to believe I have proprietary rights to discussing it. This conference signals an end to that notion.

EPIE Research Rationale

Another conference paper, the one by Dr. R. Dershimer, shows his understanding and appreciation of the work of Stake and Komoski and, as a consequence, of the EPIE Institute's approach to evaluating educational products. This approach we visualize in the manner found in Figure 1. The complete explication of the relationships within and between data-collection matrices found in Figure 1 has been detailed elsewhere (Stake, 1967b).

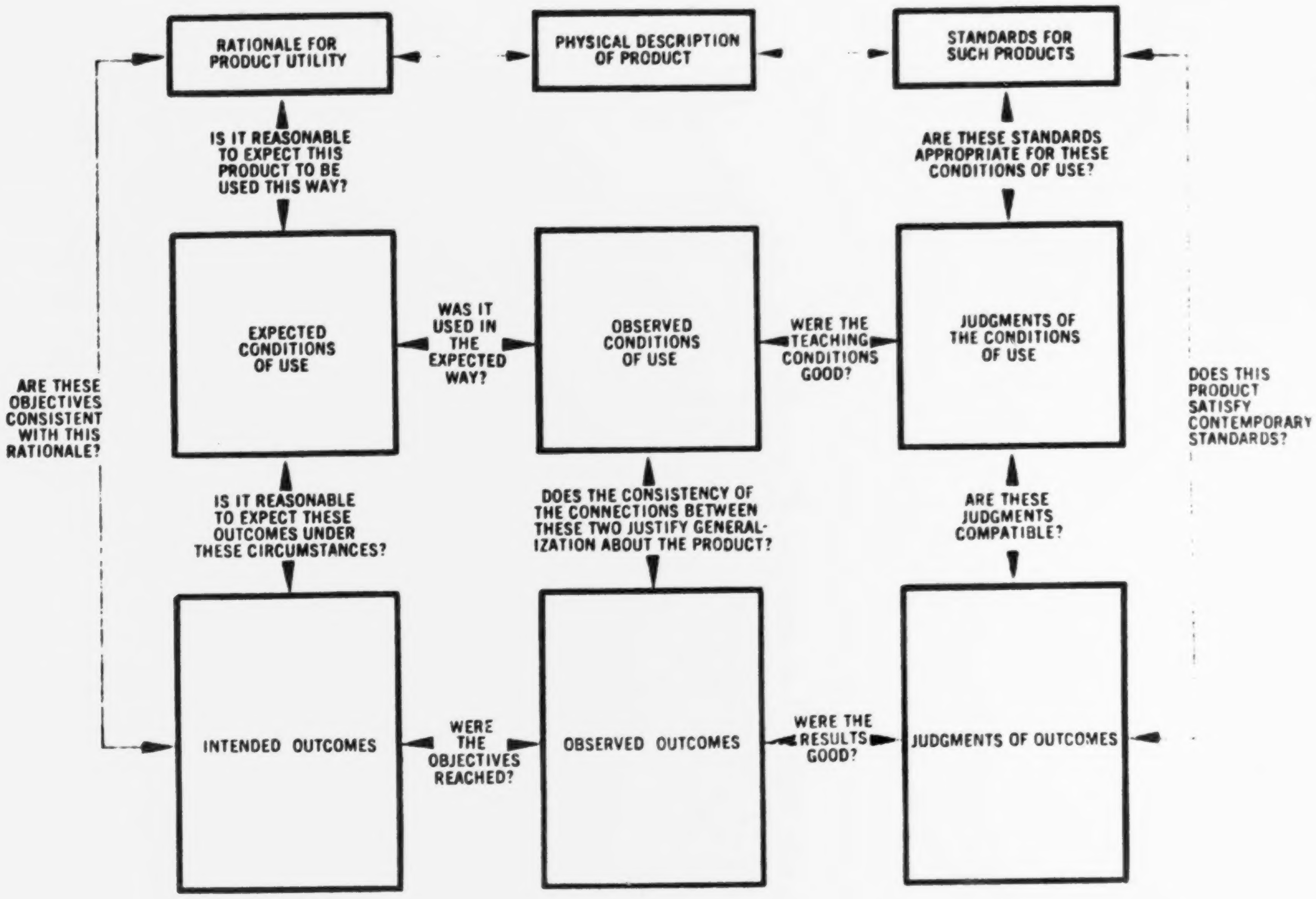
Figure 1



Matrices to guide the collection of evaluation information about an educational product.

To our way of thinking, the elements of any evaluation are bits of information. Each bit is identified according to dimensions or characteristics that help to describe the product. In the EPIE matrix designed to help the local decision-maker evaluate products, each dimension or characteristic is assigned a row. Each source of information is assigned a column. A bit of information, then, has its own sub-subcell, squared off by row and column, identified by type and source of information. We collect information from the producer, from independent analysts and from the users of the product. Next follows a representation of the processing of product information collected by EPIE.

Figure 2


RATIONALE FOR PRODUCT UTILITY
PHYSICAL DESCRIPTION OF PRODUCT
STANDARDS FOR SUCH PRODUCTS
IS IT REASONABLE TO EXPECT THIS PRODUCT TO BE USED THIS WAY?
ARE THESE STANDARDS APPROPRIATE FOR THESE CONDITIONS OF USE?
EXPECTED CONDITIONS OF USE
WAS IT USED IN THE EXPECTED WAY?
OBSERVED CONDITIONS OF USE
WERE THE TEACHING CONDITIONS GOOD?
JUDGMENTS OF THE CONDITIONS OF USE
ARE THESE OBJECTIVES CONSISTENT WITH THIS RATIONALE?
DOES THIS PRODUCT SATISFY CONTEMPORARY STANDARDS?
IS IT REASONABLE TO EXPECT THESE OUTCOMES UNDER THESE CIRCUMSTANCES?
DOES THE CONSISTENCY OF THE CONNECTIONS BETWEEN THESE TWO JUSTIFY GENERALIZATION ABOUT THE PRODUCT?
ARE THESE JUDGMENTS COMPATIBLE?
INTENDED OUTCOMES
WERE THE OBJECTIVES REACHED?
OBSERVED OUTCOMES
WERE THE RESULTS GOOD?
JUDGMENTS OF OUTCOMES

So, how will EPIE make its attack on evaluating instructional materials? Our most extensive search among the data will be for (1) congruence between what was expected of the product and what actually occurred and (2) contingency relationships between outcomes and conditions-of-use which reveal the limits of a product's effectiveness. EPIE relies on researchers and analysts with a broad range of talents and diverse methods of inference to bring about some orderly confluence of data.

Talk about Evaluation

Let us assume that the case has been made for the need to evaluate instructional materials. Let us assume also that such evaluation is wanted by several interested publics. The EPIE evaluation model is one such way of doing a job. It can be done. We hear, understandably enough, voices of vested interests who say that the educational research and evaluation community has insufficient tools at hand, few skills perfected, no strategies appropriate for beginning the job.

The catch phrases are so appealing: "Researchers can't even agree among themselves; study after study shows no significant differences; the problems under attack are trivial; the methodologies needed to do the job right aren't available." These assertions are appealing because if mistakenly elevated to the status of warranted assertion they relieve us of our obligation to evaluate: "It can't be done anyway..." "...and keeps the market the way it 'ought' to be." "The decisions should be made by the producer." That sort of thinking may have served someone well in the past. I don't know very much about how educational materials came to be marketed in the past. The little that I do know suggests that there has been considerable nonsense going on in the name of producers meeting the instructional needs of teachers. We can change the producer-controlled market of doing what comes naturally, to a user-centered market of doing what comes necessarily (Denny, 1967).

The task of evaluating instructional materials on a systematic, comprehensive scale requires a large disinterested group of competent professionals cooperating to gather, analyze, scrutinize and report on their use of educational products. It requires an enormous passel of information, reliable information, synthesized in a defensible public fashion, and a receptivity to its importance and utility by potential decision makers. We conceptualize such a strategic undertaking as shown in Figure 3. We are very excited about its possibilities, impatient with our slow progress, prayerful about necessary funding, and delighted by our reception in a few school systems with which we soon will be working closely to develop our services.

Figure 3

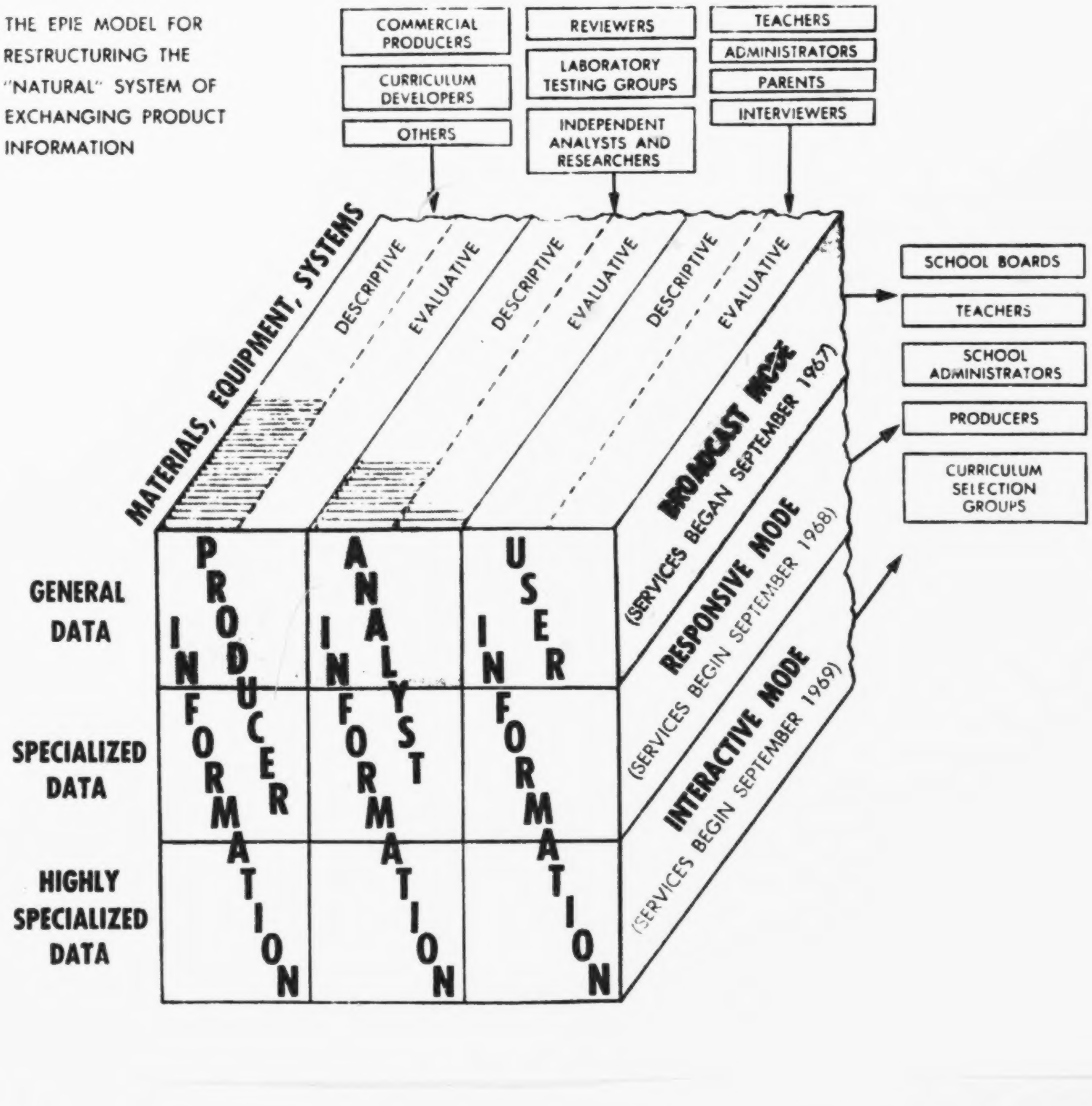

THE EPIE MODEL FOR RESTRUCTURING THE "NATURAL" SYSTEM OF EXCHANGING PRODUCT INFORMATION
COMMERCIAL PRODUCERS
CURRICULUM DEVELOPERS
OTHERS
REVIEWERS
LABORATORY TESTING GROUPS
INDEPENDENT ANALYSTS AND RESEARCHERS
TEACHERS
ADMINISTRATORS
PARENTS
INTERVIEWERS
MATERIALS, EQUIPMENT, SYSTEMS
DESCRIPTIVE
EVALUATIVE
DESCRIPTIVE
EVALUATIVE
DESCRIPTIVE
EVALUATIVE
PRODUCER INFORMATION
ANALYST INFORMATION
USER INFORMATION
GENERAL DATA
SPECIALIZED DATA
HIGHLY SPECIALIZED DATA
BROADCAST MODE (SERVICES BEGAN SEPTEMBER 1967)
RESPONSIVE MODE (SERVICES BEGIN SEPTEMBER 1968)
INTERACTIVE MODE (SERVICES BEGIN SEPTEMBER 1969)
SCHOOL BOARDS
TEACHERS
SCHOOL ADMINISTRATORS
PRODUCERS
CURRICULUM SELECTION GROUPS

Figure 3 represents the EPIE Institute's program for improving the natural system of information exchange about educational materials. The shaded areas represent information now in the system and being disseminated through our "broadcast" mode: The EPIE Forum.

The schoolman knows that out there somewhere is descriptive information about all elementary science series or all secondary school social studies materials, and about all closed-circuit TV equipment or all the overhead projectors now on the market. He also knows that in some cases analytical reviews of materials have been made by competent professionals and that laboratory tests have been conducted on some types of equipment. Further, he knows that other schoolmen have had experience with many of the products among which he must choose. Being aware that this information does exist, he does his best systematically to bring what he can together in order to make a decision. But given the number of products, their increasing complexity, and the time and staff limitations within which he works, his best efforts too frequently end in frustration and failure.

In order to carry out its objectives, the Institute must collect in a systematic fashion the information which it will process and disseminate among educational decision-makers in schools and in industry. The three columns on the face of the model refer to the three types of information (identified by the three sources, Producer, Analyst, User) that are being gathered and will eventually be disseminated by EPIE. On the left side of the model presented in Figure 3 are three levels of information sophistication: General Data, Specialized Data, and Highly Specialized Data. Down the right side of the model are represented the three basic "modes" of information services to be provided by the Institute.*

Some Problems

EPIE has been working on some problems in which I think you are interested--problems which confront you in your evaluation projects with the materials of special education. Our primary methodological problem is one of relating goals and conditions-of-use to outcomes. Other problems vie for attention: What are our target populations? To whom shall we ask our questions? What products do they want to know about? What questions aren't they asking that could be relevant to the solution of their problems? What steps must be taken to assure that classroom conditions are representative?

We have the problem of comparing products with different purposes. No two instructional aids have identical objectives. We ask some questions which show each product in its best light, as well as in light which may be best for its competitors. One product appears better under certain limited conditions, poorer elsewhere. We must aid the potential user to see what objectives, which conditions of use complement which products for his instructional needs.

---

*For a complete discussion of this model see EPIE Forum, vol. 1, No. 5, January, 1968, p. 28-33.

Still another problem has to do with standards. Most operating standards are idiosyncratic and unconscious, serving to shape personal preferences, perhaps very consistently, but avoiding public exposure. The advocate of this or that standard may adhere to still others in his own practice. EPIE's purpose is not to show what is popular, but to reveal--to expose--the various expectations that exist. Surveying every expectation is all but impossible; utterances both pertinent and suitably documented are hard to find. A thorough presentation of existing standards is a formidable obligation, but a necessary one for a nationwide evaluation project.

Identifying school goals is important, too. Are products differentially useful depending on school goals? Of course they are. Along with other statements of opinion and judgment, school goals have a translation problem. Each goal has implications for practice, but spelling out what practices are consistent and what are inconsistent is not an easy task. How can goals be quantified unambiguously? EPIE is working on the development of more definitive scales for goal priorities.

It is necessary to take the individual teacher's and the individual school's goals into account. The accounting procedure we employ compares the congruence of hopes with realizations. When they are matched, or nearly so, our report to our client might be, "At this point in our deliberations this is the match or mismatch." This is not necessarily good or bad, by the way. People can intend foolish things which fortunately go awry. One also might have quite narrow intentions for a given program which has been observed to have rather far-reaching benefits--another fortuitous mismatch.

An observed lack of congruence between intended antecedents, transactions or outcomes and observed antecedents, transactions or outcomes may be undesirable, may be merely tolerable, may be quite nice indeed, depending on the context in which the incongruity occurs. When the highly prized aims are not realized in action, it is always an unhappy experience. When unanticipated events transpire which are of some seeming potential but which could be of possible disservice to the program, the findings are indeterminate; and when outcomes exceed our more ambitious goals and are accompanied by grand benefits, the best of all possible mismatches occurs. The judgment depends on the contingencies and congruence throughout the system and not upon the presence or absence of congruence. Stake's model reflects considerable concern for the wishes of the individual instructor. But he says more.

> Consciously and unconsciously, people have different expectations of the products they use.
>
> The EPIE research rationale has no stronger commitment than the commitment to record and to honor this diversity of values. No product evaluation can be complete without a survey of the preferences and priorities of the many groups of people who use the product, or who may benefit or be injured by it.
>
> But this commitment does not preclude forthright statements of relative values (Stake, 1967b).

The congruence of local intents and observed events must be judged against external norms, general standards. Evaluation is required. These steps are required in an evaluation study but may be quite entirely absent from experimental research and assessment or status studies.

If such judgmental procedures reveal little commonality for the local school's purposes and the external criteria, difficulties are imminent. Difficulties of the sort experienced by any evaluator who tells a client that which the client would rather not hear. Difficulties of the sort which cause strained evaluator-client relationships. Hopefully evaluators can gain sufficient respect from their clientele to avoid the stigma of persona non grata. I have now come to view as vital a relatively unexplored dimension of educational evaluation, namely that of the psychodynamics of advice and information giving. We need some generous help from brethren in the business of clinical counseling.

Research or Evaluation

I would like to turn now to a few problems I have been worrying about of late regarding distinctions between evaluation and basic research, between generalizability goals and evaluation goals.

Basic research is much in the news of late. An articulate basic research proponent, Kerlinger, has written, "It just seems to be too hard to understand the nature and purpose of basic research and too easy to talk easily about applied research, research and development, product-oriented research and similar kinds of jazzy things. Besides, one gets money for the latter and not for the former" (1967, p. 5). In his acerbic commentary he charges USOE with hindering the development of basic research in education and commends Wittrock (1967) who wearies of product research "that gets us nowhere." Wittrock calls for support of basic research of educational phenomena, research for conceptual space points, for theoretical experiments in naturalistic settings. But Kerlinger regards the likelihood of getting anyone but a handful of kindred souls to listen to their pleas as unlikely. Bright's (1968) pro-basic research comments are widely printed and discussed; Cronbach's Kappan (1966) statement is quoted by everyone and now I have joined them; Gagné (1966) gags on loose molar discussion by educators about curriculum reform, evaluation, revolution; and Bruner's bombshells periodically explode about us in search of a theory of instruction. All of this could be seen as a "movement" of sorts. It is the sort of movement I am not very excited about. Basic research as I understand its potential for evaluation problems or educational decision making is not what EPIE and SEIMC really need. Basic research will of necessity take a long time to come to fruition. The vineyards need to be worked in the meantime. If not basic research, then what? More theory? More information? Yes and no. I am much taken up with the notion that we are about to be glutted with information. We are information rich. We may be headed for a period of superabundance. People are being told the information will soon be at our fingertips.

Unfortunately EPIE may be seen by some to be another member of a growing list of do-gooder agencies committed to the proposition that "what the educational world needs now is data, more data." We dislike this image very much. The educational world can do without EPIE, IBM, ERIC, SEIMC tooling up to collect, store, and retrieve still more information. I think we presume too much when we presume the usefulness of data per se. I think we presume too much when we assume the rationality of educational decision making. Dershimer has expressed a concern for the amount of folklore in our profession. Others wish to quickly supplant the folklore with empirical evidence. This could be an incorrect ploy for us to take. Kenneth Boulding (1966) suggests that empirical evidence may not always be the better evidence for a variety of human endeavors. The pluralism of the educational enterprise may need folklore, vagueness, alogical custom to provide a sort of elastic bond to keep the mélange intact.

It may be that some teachers' behavior deteriorates, when their intuition, folklore and unthinking habits are supplanted or seriously challenged by scientific information. The ingestion of information may not effect or may even produce undesirable results on educational decision making. What a frightening thing if true! I think what all this means in part is we need more data, if, and only if, we develop more and different procedures for information assimilation by our users.

I have no doubts that educators need to exchange information among themselves about instructional tools and techniques. Yet there is apprehension among educators (and certainly among producers) about organized efforts to obtain that information--and with good reason. Stake has put it, "The hazards of prejudice are not less than the hazards of ignorance."

It is obvious to the supporters of EPIE that the need for information justifies the risk of prejudice, i.e., the possibility of encouraging an occasional unwarranted innovation or maintaining some out-dated standards. The risk can be kept small; but the need for evaluative information cannot be made small, for it grows out of the imperative need for rational decision making.

What constitutes a useful distinction between evaluation and research for those concerned with evaluating instructional materials? At the core of the educational researcher's purposes is his aim to generalize his finding beyond the people and setting utilized in his research. He works hard at maximizing the likelihood of similar findings being replicated by other educational researchers doing studies of the same type with other people, other instruments, in still other ecological settings. It's a fierce task. He is deeply concerned with the questions of external validity as well as internal validity sources (Bracht and Glass, 1967). External validity sources may be grouped into two classes: first, those which deal with questions of generalizing findings to other people, or population validity; and second, those which deal with questions of generalizing findings to other settings, ecological validity. The researcher who wants to build instructional theory, to work at curriculum evolution and revolution, to standardize tests and methodologies, is necessarily primarily concerned with population-ecology validity problems.

While the educational evaluator will be aware of these matters and will do what he can to safeguard against invalidity contaminants, the crux of his concern is not that of discovery and building principles or relationships with high generalizability to other people in other settings than those under his investigation. Which is not to say that he cavalierly disregards such concerns--rather, it is a case of priorities, of alternative emphasis. His concern will be with the applicability of extant measurement techniques and regimens to the particular population and setting being evaluated. (Large-scale evaluations must of course deal with large-scale sampling procedures, data handling and statistical inference.) But the evaluator's principal charge remains one of fully describing and fully explaining the program being evaluated so as to facilitate his evaluation of matters related to that population.

The evaluator of the local school social studies curriculum is not particularly interested in whether his findings are generalizable to still other social studies instructional programs across the state or nation. Certainly the local evaluator should not be examining the behavior of a small number of people principally to acquire predictive power to make statements about a large number of other people outside his research sample. When he achieves internal validity through random sampling procedures within his available or accessible research population his validity problem is licked.

The educational researcher on the other hand must do more. He has another inferential leap to make, that of discerning the relationship of his available sample to his target population. What practical difference grows out of this distinction?

It is better for the evaluator to invest his resources in obtaining full, reliable descriptions about a restricted set of circumstances to be judged and to let uncertainty prevail about the generalizability of his findings. This latter posture is achievable only by agencies with huge capital resources, personnel and expertise necessary to assess the large target group.

EPIE will have to conduct some basic investigations and some evaluative studies. Any research study seeks generalization, but studies differ as to the level of generalization they seek. The "basic research" study in education is usually indifferent to personnel, subject matter, locality, and time. The practitioner's inquiry usually calls for minimum generalization, because a purchase to meet some given need is in the offing. But EPIE has many clients. EPIE's studies will specify the product, and search for generalization or limits related to types of pupils, teachers, schools, and so forth.

Why have I labored the point? First, because I wanted to try it out today as a possibly useful distinction for examining the tasks of evaluating educational materials and for training educational researchers and evaluators. Second, it may help those who may be using a consulting evaluator to understand better and to participate more meaningfully in evaluation projects. Third, I hope it encourages discussion and thought about the nature of evaluation, the nature of basic research and the resultant tradeoffs inherent in the conduct of an evaluation study versus a research study.

I am not certain that my notions about the differential function of generalization in experimental and comparative educational research versus situational or local evaluation studies can be or should be defended. There are some forces that could be marshalled. I have argued for theoretical research in naturalistic settings (Denny, 1967). Recall Brunswick's (1956) strong arguments for naturalistic research and consider Cattell's (1966) assertion that the progress of psychology as a science depends increasingly on nonmanipulative designs. Bracht and Glass (1967) encourage me. They claim that while basic research serves a vital function in contributing to our knowledge of the human organism such studies are not the basis for generalization to a variety of situations in which humans normally interact with their environment. Such generalizations are fraught with indeterminable risks. If carefully controlled experimental designs really do not yield generality of findings--or more so than do more local, naturalistic approaches--I feel less constrained to strive hard for more and more basic research at all costs.

Questions Are Indicators

I would like to have you consider another distinguishing characteristic of the evaluation situation as the attention we must pay to questions such as:

a. How good is this elementary science kit?
b. Can I justify keeping this literature program next year?
c. How helpful was the parental involvement in getting that program in modern math on the road last year, and should I do something similar for my modern biology curriculum revision?
d. The superintendent who asks "which program would be best for my four low-achievement schools insofar as beginning reading materials are concerned: that linguistically based series, the one that is nearly all programmed instruction, or the experimental model that's coming out of the individually prescribed instruction research shop at the university?

Now, consider some questions which are better answered by education researchers; questions of the type listed by Bracht and Glass (1967) in their treatment of external validity of the ecological variety:

a. Is the treatment equally effective with all teachers?
b. Is the treatment effect independent of the size of group?
c. Is the treatment dependent to some extent on the use of certain audio-visual aids?
d. Is the treatment effect independent of the time of day?

These are important questions. They require careful attention to safeguarding against contaminants of internal and external validity to permit the questioner to generalize to the larger target population he has in mind. They are, clearly, quite different questions from those cited as being better answered by educational evaluation.

There are many common concerns of the evaluator and researcher. Often the questions confronting educational decision makers demand approaches that are one part research to one part evaluation. Consider the social studies research project director who wants to assess the effects of value preferences held by secondary school teachers of the social studies on student concepts of the role of the social studies in upgrading American society when he asks me where to begin his evaluation, ahead are several hours of conversational give and take, coaching, structuring for both of us. But one thing we are likely to settle upon is the need for looking at the interaction of teaching style with personological variables of the learner. We will be concerned with designing that particular facet of the study to reveal the presence of ordinal or disordinal interactions of the sort that Cronbach (1957) called for over ten years ago. Hence, this facet of the evaluation project is likely to be traditional educational research methodology pure and simple, and there will be still other research elements to be included. But when the final selection of what we shall or shall not do comes about, I will be willing to trade off maximal generalizability (or concern for sampling procedures to maximize external validity) in order to complete my picture of the local social studies instructional milieu under investigation. Of course I shall exercise extreme caution to control internal validity, as must the program researcher or the experimental researcher. But there are distinctions to be made.

I have a feeling that there is too much high-powered research motion going on in the name of evaluation. Conversely, in the worst sense of the term, that of defining evaluation as the act of imputing personal preference to matters of choice, there may be too much evaluation going on in the name of research in our schools. But let me end this tangential probe by saying there's not enough research or evaluation going on in the name of anything. The documentation for this assertion has been made by others, elsewhere, to my satisfaction, at least. (Scriven, 1967; Kerlinger, 1967; Stufflebeam, 1966; Cronbach, 1966; Stake, 1968)

Obstacles

There are school people deeply concerned by educational researchers' unwillingness or inability to listen to the real questions coming from the practitioners--the kinds of questions which should interest educational evaluators, which very likely turn off educational researchers, and . . . alas . . . are being turned down for supportive consideration by a fair number of us as impossible for study. These messy matters are the stuff on which decisions must be made. Some of us are taking all this quite seriously. Stufflebeam (1966) and Hammond (1967) have been listening to such questions, to such cries of the anguished educational decision maker, and have been developing new paradigms for coping with the content and contextual realities of evaluation problems. The research and theory of Guba (1965) encourages me to believe that the tried and untrue old tribal research dance, the pre- and post-test shuffle, can be challenged, improved upon, and where need be, discarded. The critical work of Hastings (1966), Glass (1967, 1968), Stake (1967a) and Webb (1966) and his associates support EPIE's contention that help is coming. Not tomorrow. And if gigantic financial support agencies don't help, not for many days after tomorrow either.

Evaluation is a bad word in several kingdoms. If the word "money" can be translated into "interest", "affection" or "confidence", please check "none of the above" as your answer to the question, "What fascination, love or faith do the money givers have for educational evaluation and evaluators?" Clearly they misunderstand our message; or understand and disvalue it; or understand, value and are about ready to give us a chance to develop and practice our instruments and techniques to increase our usefulness to those who want to evaluate instructional materials.

Could it be that this evaluation conference is the first trickle of a new wave of financial support about to engulf the educational community for the exploration and conduct of evaluation activities? I'm about ready for a good swim.

REFERENCES

Boulding, K. The impact of the social sciences. New Brunswick: Rutgers University Press, 1966.

Bracht, G.H., and Glass, G.V. The external validity of experiments: Research paper No. 3, Laboratory of Educational Research, University of Colorado, 1967.

Bright, R.L. Bright seeks 'basic' educational research. Educational Researcher, 1968, 8(2), 1-2.

Brunswik, E. Perception and the representative design of psychological experiments. Berkeley: University of California Press, 1956.

Cattell, R.B. Guest editorial: Multivariate behavioral research and the integrative challenge. Multivariate Behavioral Research, 1966, 1, 4-23.

Cronbach, L.J. The role of the university in improving education. Phi Delta Kappan, 1966, 539-545.

Cronbach, L.J. The two disciplines of scientific psychology. The American Psychologist, 1957, 12, 1957, 671-684.

Denny, T. Educational product evaluation. First Annual Project ARISTOTLE Proceedings, 1967, in press.

Gagné, R.M. Elementary science: A new scheme of instruction. Science, 1966. 151(3706), 49-53.

Glass, G.V., Editor. Research Papers, Laboratory of Educational Research. Boulder: University of Colorado, 1967-date.

Guba, E.G. Methodological strategies for educational change. Paper presented to the Conference on Strategies for Educational Change, Washington, D.C., November, 1965.

Hammond, R.L. A design for local evaluation. EPIC Forum, 1967, 1, 3-6.

Hastings, J.T. Curriculum evaluation: The why of the outcomes. Journal of Educational Measurement, 1966, 3, 27-32.

Kerlinger, F.N. Letter to the editor, Educational Psychologist, 1967, 5(1) (7).

Komoski, P.K. An invitation. EPIE Forum, 1967, 1 (1), 2-3.

Scriven, M. The methodology of evaluation. AERA Monograph Series of Curriculum Evaluation, No. 1. Chicago: Rand McNally, 1967, 39-89.

Stake, R.E. The countenance of educational evaluation. Teachers College Record, Vol. 68, No. 7, April 1967a, 523-40.

Stake, R.E. A research rationale for EPIE. EPIE Forum, 1967b, 1, (1), 7-15.

Stake, R.E. Testing in the evaluation of curriculum development. Review of Educational Research, 1968, 38 (1).

Stufflebeam, D.L. A depth study of the evaluation requirement. Theory into Practice, 1966, 5, 121-133.

Webb, E.J. et. al. Unobtrusive measures: Nonreactive research in the social sciences. Chicago: Rand McNally, 1966.

Wittrock, M.C. Paradigms in research on instruction. Educational Psychologist, 1967, (1) 1-3.

Evaluation: Procedures

THE EDUCATIONAL MODULATION CENTER

by Gary Adamson, Ed.D., Director
Educational Modulation Center, Olathe, Kansas

The Educational Modulation Center (EMC) concept is relatively unknown to the educational field. It is based on the premise that effective procedures can be carried out to remediate learning problems in children. Its primary emphasis is on the children since it is felt that this is the only feasible way to contact the source needed for improvement of education--the classroom teacher. When teachers have a desire to assist students with specific learning deficits, they may be instructed in clinical appraisal and systematic methods for remediation. They can also be introduced to new techniques and materials. Therefore, by working with the child and teacher simultaneously to effect a better program, future educational programs and teacher performance can be improved.

There are two unique aspects to the Educational Modulation Center: the newly designed educational team and the materials resource center. The educational team, composed primarily of educational members utilizing other services and specialists as resources, was designed because: (1) In the past, school psychologists have been the primary referral source and final determiner of educational services to teachers. This is considered to be a partially if not a totally inappropriate activity. (2) It is more appropriate to provide less expensive but more knowledgeable teaching techniques and materials before referring to a psychologist or other specialist for long-term, in-depth evaluation and diagnosis.

The materials resource center is a laboratory where materials and programs are developed for remediation and where materials are analyzed for specific remediatory functions.

The Educational Modulation Center is an Elementary and Secondary Education Act (ESEA) Title III Project and is a result of a planning grant in the year 1966-1967. The project has been funded for three years through the Office of Education, Title III, to service 24 school districts in a three county area in Kansas. The most expedient description of the project is by delineating its operational procedures.

Operation of the Program

Referrals are made by teachers and cosigned by principals, although in certain instances different channels have been established by school systems. Since the basic concern is to provide services to children and to help teachers, few referrals are initiated from outside the school district. When a referral is received by the Center, it is reviewed by the Director of EMC and appropriate suggestions for its handling are made. These include the following:

(1) Referral to Team Coordinator for processing through the educational team.

(2) Referral for psychological evaluation in such cases where there is an identification need for special education.

(3) Reject the referral because it is inappropriate.

Because of the limitations of personnel, the EMC has specified the type of child to be worked with, but this could be readily expanded to all areas of exceptionality with additional staff. We have identified what we call the forgotten child. This is the child having average or above average intelligence but who is educationally retarded by one or more grade levels in one or more subject areas, i.e. learning disabilities.

Educational Team Procedures

Approximately 80% of the referrals are handled by method and materials consultant/teachers thereby allowing time for allied members of the team to be more productive. In the past school psychologists spent much of their time giving evaluations to all referred children and by doing so greatly hampered their effectiveness as diagnosticians. Added time allows for a more appropriate role for the school psychologist in intensive diagnosis, research, consultation and counseling to parents and children. When receiving a referral a methods and materials consultant/teacher will visit the school. These visitations have many purposes but primarily involve the teacher and the consultant in an attempt to find immediate answers to the child's problems through methods and materials, arrangement for specialized services, behavior modification or other appropriate activities. Another purpose is to study the child in-depth, as well as the school to determine appropriate placements, resources and suggestions available at the school. Through planning materials and explaining these to the teacher and by showing the teacher how to manage this individualized program, the teacher must indeed learn. Through this procedure, an individual or one-to-one in-service program is carried out effectively. Since the child should not be removed from the classroom until all else fails, it is important that the methods and materials consultant/teacher become quite familiar with the school and principal as well as the child.

The methods and materials consultant/teachers study the child intensively and try to effect educational remediation by using the methods and materials and the knowledge of both the teacher and the methods and materials consultant. If there are any factors affecting the child and his learning that are beyond the scope of these individuals, referrals may be made within the team for intensive psychological, audiological and family background workups. The methods and materials consultant working with the teacher can either effect the procedures suggested by these specialists or determine through informal meetings with them other appropriate referrals to further sources such as neurological or physical examinations.

If, after all these evaluations have been made and little or no improvement has been noted, a referral can be made to a diagnostic classroom. This is a classroom used for intensive microscopic observation by methods and materials consultant/teachers, psychologists, all other members of the EMC staff, university staff, etc. By experimentation with different kinds of approaches with different materials, by using all kinds of media, therapists and curriculum specialists, a child is provided in most cases with a program which will adequately remediate his problem in the regular classroom. In certain situations a direct referral will be made to special education classes or to a residential treatment center if community services are lacking.

When the staff and diagnostic teacher are satisfied that a beneficial program has been planned for the child, he is then returned to his regular classroom. Before his return, the classroom teacher visits the diagnostic classroom for observation of the child and is given a "packet" that has been prepared by the team member consisting of lesson plans and materials for the child upon his return. Continuous service by the methods and materials consultant/teacher is provided during the initial period of return; a periodic follow-up service is also carried out.

Toward the end of the year, the child will be re-evaluated and his individual program re-examined. If the program has not been as successful as anticipated, then the child will attend a summer session to be taught by methods and materials consultants, diagnostic teachers and psychologists. (Let's keep specialists honest.) In this summer diagnostic setting, the child will receive intensive remediation and re-evaluation and his new teacher for the next year will be brought in before school starts for an explanation of the materials and procedures planned for the child. Consequently, this provides for continuous follow-up, re-evaluation and non-termination of difficult cases. If at any time during the next year, the child seems not to be progressing with the planned program, he can of course be re-admitted to the diagnostic classroom, for further experimentation and program planning.

A few other important aspects should be mentioned which culminate and solidify the functioning of the team. The first of these is the materials laboratory which is based on the premise that learning materials can be adequately broken down and analyzed to be used effectively with various kinds of learning problems. In this way, specific materials may be recommended to deal with a specific problem after it has been diagnosed. The materials laboratory at the EMC is divided into many areas of curriculum. Reading, for example, can be broken down into several sub-categories, and these sub-categories, such as phonics, can in turn be divided into even smaller units. These specifications must be carried out further, making provisions for mental age, grade level or reading level equivalents as well as for format (i.e. basal reader, transparency, programmed material, etc.). The Modulation Center has developed a procedure whereby appropriate materials can be quickly selected after diagnosis and be made available to the child immediately.

Each piece of material received by the center is examined in detail. This is an objective examination accomplished by actually going through the material and breaking it down into component parts. The specific information obtained from this analysis is then coded and becomes the selective factor in the analysis system. The specific analysis is supported by a narrative section providing further general information concerning each skill covered and descriptive characteristics of the material. All cards in the analysis system are set up for manual selection by key sort.

After a teacher has used material from the lab, she is requested to fill out an evaluation card which is filed for the information of the EMC staff and others who might possibly use the material. A detailed description of evaluation and analysis procedures, impossible in this brief outline, may be obtained by writing the center.

A second important aspect is in-service training for teachers. Teachers can become knowledgeable concerning methods and materials through one-to-one in-service instruction. But greater discovery and experimentation must be provided and this can only be accomplished through training in large numbers. In-service programs designed to instruct the teacher in educational management, application and location of new methods and materials including machines, media and allied resources should all be provided.

A third element concerns research in the classroom. We have known for some time that teachers receive very little of the reinforcement that they deserve. Curriculum directors, college professors, research directors and superintendents have taken all the credit. Teachers have always been trying methods and techniques and utilizing materials effectively. Although they have been doing these things in their classrooms, Polly down the hall has not known what Sally was doing because Sally either did not know how or was too modest to research them and disseminate the results.

Through this practical procedure called the Education Modulation Center teachers can become clinical teachers, children having educational problems can be helped and new methods and materials can be developed at the local school level--yes, even at the classroom teacher level where it should be done.

For a long period of time, we have had at our disposal the ability to diagnose a child's learning problem. This diagnosis has been done primarily by school psychologists, and occasionally by classroom teachers, reading consultants, and curriculum directors. Generally speaking, the school psychologists gave an intellectual measure, a perceptual measure, and some individual standardized achievement tests. The reading teacher generally looked at the child's reading and how well he was comprehending, his reading rate and ability for word attack. The curriculum director, using the standardized intelligence tests, suggested that the child needed help in word study skills, language arts, artithmetic, etc. Although all of these methods lack the precision needed for the adequate diagnosis of educational problems, they have existed for a number of years and have become the accepted way of evaluation.

Materials analysis is a process that has also existed for some length of time. Materials have been examined to see how they might best be used, i.e., are they appropriate for teaching a child phonics or for teaching a child about South America? Although book companies, authors and other individuals have been asked to analyze materials for their usefulness, the result has usually been a gross analysis saying that this material is good for phonetic training or good for this or that. We have never had specific analysis such as, "this material is suitable for the training of beginning sounds from pages 1-12 and ending sounds, page 14-16. This material teaches through the method of overlearning and is available in programmed form."

Diagnosis is not new to the educational field and materials analysis has been going on for several hundred years. What has been lacking is a way of putting the two together. The key, then becomes how does one, based on a diagnosis, prescribe individually and precisely to remediate a child's learning problem.

The EMC has taken ideas for retrieval of materials after they have been analyzed from other organizations. Science Research Associates (SRA), for example, uses a pick system as a way for their salesmen to select cards for general titles and subject areas. We have taken this idea but refined it for our own purposes so that the cards are broken down into very precise areas.

The general area of phonics for example, is broken down into diphthongs, blends, etc. This system allows for the immediate retrieval of a material after it has been analyzed.

However, after a material has been analyzed specifically and is easily retrievable, other areas must still be investigated. A material must be field-tested to see if it really does what it is supposed to do in terms of remediating specific disabilities.

To make the diagnosis of a child's problem more precise, the Modulation Center is compiling a "behavioral expectancy learning profile". A diagnostic instrument such as this is constructed by recording the child's behavior. Can he add 1 and 1? Is he able to add three place numbers, to carry? Can he do beginning sounds, ending sounds? On what level does he comprehend? How does he best comprehend, visually, auditorially or both? If such information along with the child's mental age can be accurately recorded, then by retrieving previously analyzed and evaluated materials that are appropriate and specific to his problem, the child receives an educational prescription.

After this has been accomplished, it is still necessary to ascertain whether or not the material works successfully with the child. If it proves to be successful, we have gained a record of how to use a specific material. If a material does not work, a new program must be devised for the child. If it proves to be successful, we have gained a record of how to use a specific material. If a material does not work, a new program must be devised for the child. This is perhaps the most important aspect. It is time in the field of education that we become secure enough to say we are wrong; we need to realize that there are other methods and programs for remediating a child's disability and be willing to use these.

The concern then becomes how one is able to test the efficiency of this procedure over the normal procedure of diagnosis, then hunting for some way to remediate the child's problems by using a variety of materials to determine what is appropriate. We have established a procedure which when finally perfected will allow us to evaluate the proficiency of fitting materials and diagnosis together.

Following the initial diagnosis, all diagnostic information is prepared for data processing. The data is then analyzed via cluster analysis. In this manner children with similar diagnostic profiles are clustered together. The clustering of children with similar diagnostic profiles constitutes the diagnosis of the specific type of learning difficulty. It is recognized that this procedure which is essentially a statistical diagnosis of a specific learning disability, constitutes one way of defining learning disability and at the same time considers numerous variables that may be associated with learning problems. The evaluation of remedial materials procedures is based on the specific type of learning disability as reflected by the profile analysis. The profile analysis is the cluster analysis that is gained from the diagnostic initial information. In this manner, the material and/or procedures is based on the specific type of learning disability as reflected by the profile analysis. The profile analysis is the cluster analysis that is gained from the diagnostic initial information. In this manner, the material and/or procedure for a given child and a given cluster is evaluated as to appropriateness and applicability. For example, a given piece of material may be used with children in several clusters. After a given period of time the children are re-evaluated to determine if remediation has been successful. In other words, based on diagnosis, materials were prescribed, these materials must now be evaluated to check the validity of the prescription for the diagnosis. It would be quite possible that children in one cluster may make somewhat greater gains than children in other clusters. If so, this would give some indication as to the appropriateness of the material for the various clusters. This does not imply that on the basis of one trial all future children in the successful clusters will automatically be given the same material, but instead it does mean that a given material could be re-evaluated under markedly more rigid conditions. This indicates that if a specific material has been tried with success then the likelihood of it being successful again is probable. In summary, the process of diagnosis and evaluation of materials and all procedures may be based on the following steps:

1. The child is evaluated by the educational team.
2. After the initial diagnosis of learning disability is obtained, specific diagnosis of type of disability is established via cluster analysis placing children with similar disabilities together.
3. Evaluation of remedial procedures and the use of materials is based on the appropriate method and applicability of cluster.

Even though diagnosis and materials evaluation have been available for a long period of time, few people except in a few scattered and isolated situations have attempted to fit diagnosis of educational problems and prescription of materials together to test the appropriateness or inappropriateness in prescribing for children. Additional information may be obtained on cluster analysis by writing Dr. Gary Hoeltke, c/o The Educational Modulation Center, 310 N. Marion, Olathe, Kansas 66061 or Dr. Gary Adamson at the same address for any aspect of the EMC.

THE EVALUATION OF INSTRUCTIONAL MATERIALS

by Floyd G. Hudson, Ed.D.
Assistant Professor, School of Education, University of Kansas
Associate Director, Educational Modulation Center, Olathe, Kansas

Based on past experience and observation we may conclude that educators, in general, are guilty of selecting instructional materials based on "hearsay", "intuition", "statements of the publishing company representative", "pictures in a catalog", display of materials at a favorite convention", etc., etc. Conclusions may be drawn to infinity, but this does not change the behavior.

If educators were confronted with unscientific selection and use of medicine by the medical profession, we would be quick to encourage the enactment of control measures, either from the governmental level or from within the profession, yet as educators we perpetuate an act not allowed to others.

It is therefore imperative that as professional educators we serve as safe-guards to the education profession. This may entail obtaining assistance from other sources such as central and/or regional centers serving as clearing houses for the collection and dissemination of information relative to instructional materials.

The emphasis at the Educational Modulation Center (EMC) in prescriptive programming for children with learning problems necessitated the development of a process through which the evaluation of instructional materials may take place.

The systems of Materials Analysis: Criteria Establishment, and Evaluation of Instructional Materials are based on current needs.

We have established a framework and are in an early stage of development. Our functional importance is our level of practice: the level of the child within the classroom.

We try not to disappoint our colleagues, as we have not coined a new "verbiage" nor have we cast aside ideas from yesteryear, nor have we copyrighted a new "gimmick" in order to obtain federal funding.

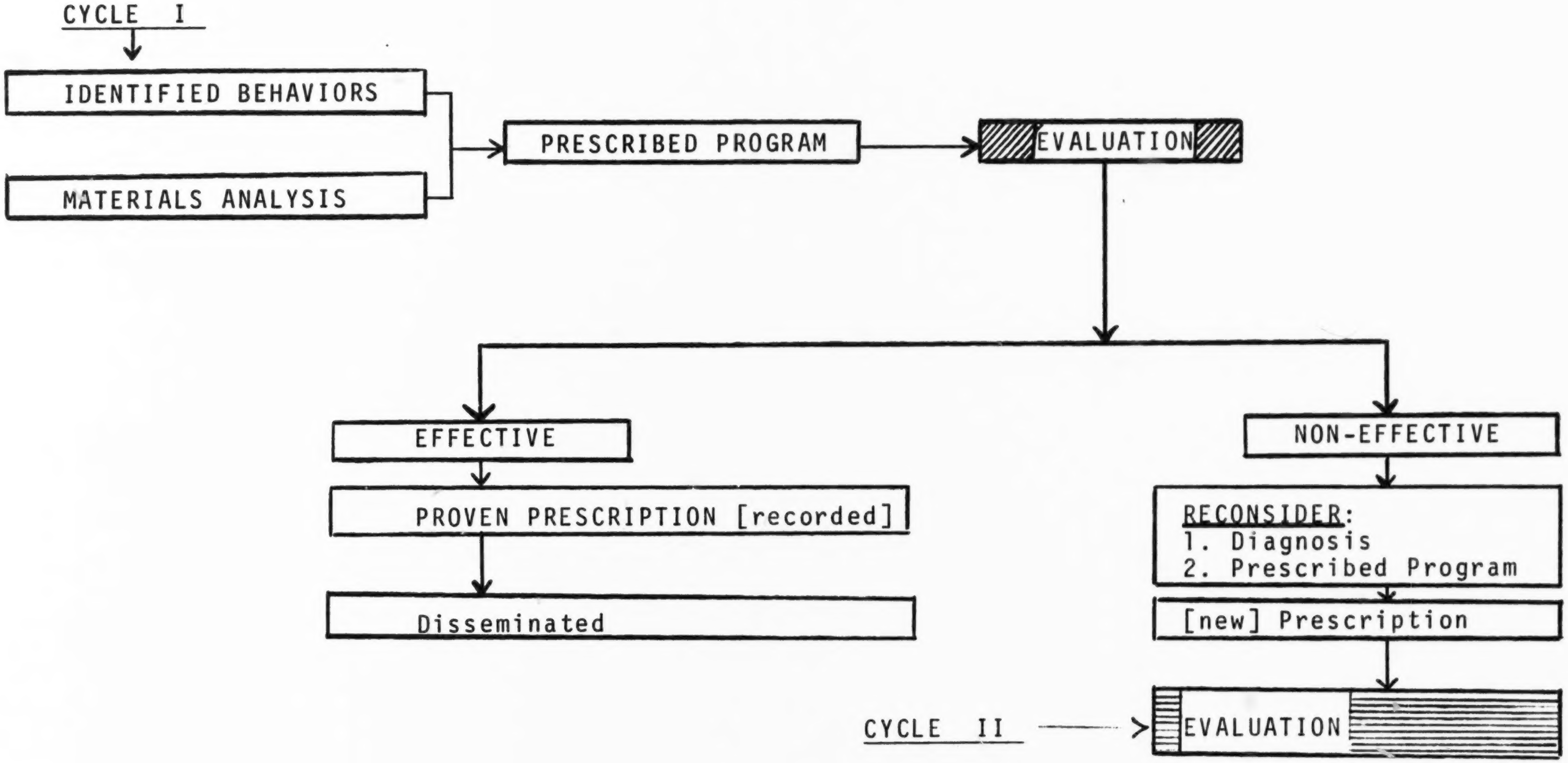
We are confident that in order to improve the functioning of our materials lab resources and to develop a useable prescriptive program, we must effect and validate growth and change in the behavior of children.

Instructional Materials Evaluation System

The Instructional Materials Evaluation System is best illustrated by diagram #1. The diagnosis and prescriptive program are dependent upon an analysis of the (+) positive and (-) negative learning behaviors of the child [IDENTIFIED BEHAVIORS] as well as analysis of instructional materials [MATERIALS ANALYSIS]. In order to prescribe a program for the child with learning problems

DIAGRAM #1

E M C INSTRUCTIONAL MATERIALS EVALUATION DIAGRAM



the materials must be analyzed as to their component parts and function.

It is important to keep in mind that the analysis of materials, in and of itself, does not substantiate the effectiveness purported by the author or publisher. It is only from evaluation that this can be done.

Once this analysis of learning problems and materials is completed, the diagnostician is concerned with the matching of one to the other. A gross example may be an identified weakness in phonetic skills and the selection of an instructional material stimulus, designed to elicit a response, which will provide the desired behavioral change.

As can be readily ascertained, the evaluation of instructional materials has as its initial focus the problem centered on behavioral approach. The identification of behavioral symptoms is done by the educational team (composed primarily of educational members utilizing other specialists and services as resources) with the primary initial responsibility being held by the methods and materials consultant/teacher (M & M). Information relative to diagnosis and prescriptive programming is recorded on Contact Report Forms, diagram #2, each time a member of the educational team is in contact with the referred child. The information recorded serves as valuable material in the evaluation of the prescribed program of which materials are the most important element.

The evaluation of instructional materials must have as its basic criterion from which we establish hypotheses and make certain assumptions relative to the effectiveness or noneffectiveness of the material to do that task for which it is intended. A system for materials analysis, which in effect establishes a criterion for judging the merit of the material was developed early in the project year by this writer and was expanded during the initial phase of the project.

Instructional Materials Analysis: Criteria Establishment

The criteria for judging the merit of those materials are based on objective data; therefore we make our basic assumptions based on this kind of information as presented by the author and/or publisher. If we take the published data and make an interpretation of the material as to what we think it will do, we have then taken our criteria establishment out of objectivity and have both objective and subjective judgments, which is in effect a form of subjective evaluation. A Descriptive Analysis Worksheet has been developed for recording the descriptive information relative to a particular instructional material. Diagram #3 illustrates the format to be followed in recording this information. Diagrams 4, 5 and 6 illustrate completed analysis worksheets. (The fourth diagram is a descriptive analysis of a series; #5 and #6 are descriptive analyses of the individual components of the series.)

Once a descriptive analysis has been completed this information is recorded onto the Descriptive Analysis Card, with discrimination areas designated for easy retrieval and for the matching of materials to learning problems. This analysis makes the usefulness of the materials readily apparent to the person desiring a particular instructional material.

DIAGRAM #2

Referral No.____________

EDUCATIONAL MODULATION CENTER

CONTACT REPORT NO._________

Student:______________________M F Grade:__________________ Date:__________

Teacher:_______________________School:__________________Dist:_____Area:___

---

(1) Conferences: Teacher ( ) Admin. ( ) Parent ( ) Other:__________________

(2) Student Observation:

(3) Diagnostic Evaluation: (test(s))

Observations:

(4) Remediation: (Materials)

(5) Remediation: (Techniques)

(6) Summary Notes:

(7) Dispensation: Contact Continuing__________ Contact Terminated____________

(8) Contact Schedule: ____

(9) Use Reverse Side: (Yes) (No)

____

Methods and Materials Consultant

EMC/67/600

The illustrations used here deal specifically with one particular instructional area (READING). Diagram #7 illustrates the Materials Analysis Guide Card: Reading. Diagram #8 illustrates the Problem and Format Card: Reading (Phonics). Diagrams #9, 10 and 11 illustrate the recorded Descriptive Analysis Card: Reading (Phonics), for the series as well as each of the component parts of the series. It becomes readily apparent to the teacher the usefulness of such a system as individual programs are being developed for children. Cards are prepared for each piece of instructional material and placed into corresponding files for easy retrieval. As we become larger in volume, the system should be placed into a computerized program for easy retrieval.

Instructional Materials Evaluation: Types

The evaluation of instructional materials at the EMC can best be described by this illustration denoting the three major forms. (Diagram #12) The selection of a specific type will depend not upon the needs and personnel to carry out this function; but upon the particular materials selected for evaluation.

The first type is that of Opinion. This form of evaluation is subjective in nature, although it is relatively inexpensive and more expedient.

The format developed for Opinion evaluation is partially structured in order to reduce the subjective nature of the results. Reliability becomes apparent by replication of this form of evaluation by each teacher who uses it, thereby developing more objective evaluation results. Diagram #13 illustrates the Materials Evaluation Form currently in use at the EMC.

The Opinion Evaluation is obtained from these possible sources:

(1) Classroom Teachers: The instructional materials at the EMC are available to teachers on a loan basis for use in the classroom and/or any other professional reason. Once a teacher has had this opportunity we ask that they evaluate the material (see Diagram #13).

(2) EMC Staff: Each member of the Educational Team charged with the responsibility of prescriptive programming for children completes an evaluation of materials form based on the results from their use in the child's program.

(3) Interns: Students assigned to the EMC as a part of their practicum requirements are asked to complete evaluations based on their professional interpretation and any special uses made of the material during their internship.

(4) Authors and/or Publishers: This information is derived from the promotional literature.

(5) Other: This might well be a committee or panel of professional personnel appointed specifically for this purpose.

Quantification Evaluation is the second type of evaluation carried out at the EMC. Quantification of specific characteristics of instructional materials is for the most part objective but may become subjective depending upon the

DIAGRAM #3

DESCRIPTIVE ANALYSIS CARD

(worksheet)

SHELF NO.____________________ READING LEVEL (Publishers)____________

FORMAT NO.____________________ READABILITY CHECK________________

PROBLEM NO.____________________ MENTAL AGE____________________

MAJOR AREA(s)____

SUB-CATEGORY____

TITLE____

Titles within series: (List reading levels beside titles)

AUTHOR____

PUBLISHER____

COPYRIGHT____________________

SPECIFIC CONTENT ANALYSIS:

DESCRIPTIVE ANALYSIS SECTION:

1. General statement of what author or publisher says material is designed to do.

2. Amount of coverage on various skills taught. Be as exact as possible.

3. How does the author present his lessons or material.

4. How much teacher time is involved?

5. Size of group for optimal use

6. Time element

7. Manipulation of materials

8. Accompanying Apparatus

9. Illustrations

10. Paper and Printing

11. Manual

12. Suggestions

13. Correlated materials

DESCRIPTIVE ANALYSIS CARD

(worksheet)

SHELF NO. PH1-PH1.5

READING LEVEL (Publishers) Phonically Graded

FORMAT NO. A, B

READABILITY CHECK

PROBLEM NO. 18

MENTAL AGE 6.5-9.5

MAJOR AREA(s) Reading

SUB-CATEGORY Phonics

TITLE The Easy Road to Reading Improvement (Series)

Titles within series: (List reading levels beside titles)

Pat, Lad and the Sleepy Pig
Rednose the Elf

AUTHOR Guilano, William

PUBLISHER Marand Publishing Company

COPYRIGHT 1960

SPECIFIC CONTENT ANALYSIS:

18-Series

DESCRIPTIVE ANALYSIS SECTION:

1. General statement of what author or publisher says material is designed to do.

The Easy Road to Reading Improvement consists of two readers with accompanying workbooks. Words containing consonants and long vowels are used predominantly in the first stories. A few non-phonetic words are presented. The author intends his material to be used with elementary school age children who wish to improve their reading skills.

2. Amount of coverage on various skills taught. Be as exact as possible.

Words containing new phonic elements are introduced every five to fourteen pages. Each of the 461 basic words is used on at least 8 different pages of the reader and on at least five pages of the corresponding workbook.

3. How does the author present his lessons or material.

Workbooks accompanying the readers contain the same basic words and phonic elements in corresponding units. The child first does a unit in the workbook before he reads the corresponding unit in the reader. New words are first taught as sight words. Through the use of games and exercises the words are then analyzed phonically.

4. How much teacher time is involved?

This material does require close teacher direction. A large percent of the material requires active participation of the teacher. Some workbook pages can be completed independently of the teacher after proper instruction.

5. Size of group for optimal use

This material lends itself particularly well to individual or small group work, although it could be used with large groups.

6. Time element

It is easily used in the time usually allowed for a reading program or in even smaller time blocks.

7. Manipulation of materials

8. Accompanying Apparatus

9. Illustrations

Illustrations in the reader are colored, simple drawings with little detail. They appear either at the top or bottom of the page to allow for an even flow of print across the page. Workbook illustrations are simple sketches making use of a small amount of neutral colors.

10. Paper and Printing

The paper is a non-gloss with large print.

11. Manual

There are two separate manuals. One is written for educators and a less complicated edition is provided for parents and laymen.

12. Suggestions

Since this series relies heavily on the skills of structure and configuration it would probably be best used with children who have good visual retention and recall. Children having difficulty in this area would possibly have their reading problem complicated through initial use of this series in remediation. Because of the use of the basic Dolch sight words, books utilizing these words in abundance would be good follow-up materials.

13. Correlated materials

Consonant Pictures for Peg Boards

DIAGRAM #5

DESCRIPTIVE ANALYSIS CARD
(worksheet)

SHELF NO. PH1.2 - PH1.3

READING LEVEL (Publishers) Phonically Graded

FORMAT NO. A, B

READABILITY CHECK

PROBLEM NO. 2,3,4,6,7,8,9, 14,15,20

MENTAL AGE 6.5 - 9.5

MAJOR AREA(s) Reading

SUB-CATEGORY Phonics

TITLE Rednose the Elf

Titles within series: (List reading levels beside titles)

Pat, Lad and the Sleepy Pig, (Easy Road to Reading Improvement Series)

AUTHOR Guiliano, William

PUBLISHER Marand Publishing Company

COPYRIGHT 1960

SPECIFIC CONTENT ANALYSIS:

2 - Initial Consonant
3 - Long Vowel
4 - Short Vowel
6 - Digraphs
7 - Diphthongs
8 - Endings
9 - Blends
12 - Silent Letters
20 - Rhyming

DESCRIPTIVE ANALYSIS SECTION:

1. General statement of what author or publisher says material is designed to do.

The author intends his material to be used with elementary school age children who wish to improve their reading skills. Rednose the Elf presents 279 new words beyond those used in the A reader of this set.

2. Amount of coverage on various skills taught. Be as exact as possible.

Four pages of U.(1) deal with long and short vowels. Two pages on rhyming. The digraph th is also introduced. In U.(2) the letter j is introduced for the first time. About a third of the unit is diphthong work, the rest digraphs and blends. U.(3) is on blends, endings and more complicated work in approximately equal number. In U.(4) the primary emphasis is on blends, although much review material is included. All work in unit five is a review of previous work except for the presentation of x. Most review is concerned with word families. Unit (6) also follows a review pattern but introduces the letter z and the ook sound. Unit (7) presents the qu and deals with ir, er, and ur sounds. The new work in Unit (8) deals with or, au, ew sounds on four pages. Each unit contains an Orphan and Lost Words section (non-phonetic words) and a Let's Have Fun section which utilizes new learnings.

3. How does the author present his lessons or material.

Workbooks accompanying the readers contain the same basic words and phonic elements in corresponding units. The child first does a unit in the workbook before he reads the corresponding unit in the reader. New words are first taught as sight words. Through use of games and exercises the words are then analyzed phonically.

4. How much teacher time is involved?

This material does require close teacher direction. A large percent of the material requires active participation of the teacher. Some workbook pages can be completed independently of the teacher after proper instruction.

5. Size of group for optimal use

This material lends itself particularly well to individual or small group work, although it could be used with large groups.

6. Time element

It is easily used in the time usually allowed for a reading program or in even smaller time blocks.

7. Manipulation of materials

8. Accompanying Apparatus

9. Illustrations

Illustrations in the reader are colored, simple drawings with little detail. They appear either at the top or bottom of the page to allow for an even flow of print across the page. Workbook illustrations are simple sketches making use of a small amount of neutral colors.

10. Paper and Printing

The paper is a non-gloss with large print.

11. Manual

There are two separate manuals. One is written for educators and a less complicated edition is provided for parents and laymen.

12. Suggestions

Since this series relies heavily on the skills of structure and configuration it would probably be best used with children who have good visual retention and recall.

13. Correlated materials

Consonant Pictures for Peg Boards

DIAGRAM #6

DESCRIPTIVE ANALYSIS CARD
(worksheet)

SHELF NO. PH1.1 - PH1.2

READING LEVEL (Publishers) Phonically Graded

FORMAT NO. A, B

READABILITY CHECK

PROBLEM NO. 2,8,9,.2,.9,20,21

MENTAL AGE 6.5 - 9.5

MAJOR AREA(s) Reading

SUB-CATEGORY Phonics

TITLE Pat, Lad and the Sleepy Pig

Titles within series: (List reading levels beside titles)

Rednose the Elf (Easy Road to Reading Series)

AUTHOR Guiliano, William

PUBLISHER Marand Publishing Company

COPYRIGHT 1960

SPECIFIC CONTENT ANALYSIS:

Initial C. -2
Endings - 8
Blends - 9
Silent Letters - 12
Discrimination - 19
Rhyming - 20
Sound - Symbol Matching 21

DESCRIPTIVE ANALYSIS SECTION:

1. General statement of what author or publisher says material is designed to do.

The author designed his material to be used with the elementary school age child who needs to improve his reading skills.

2. Amount of coverage on various skills taught. Be as exact as possible.

Pat, Lad and the Sleepy Pig reader contains 182 basic words. The workbook uses some extra words which are almost exactly like the basic words. The workbook is divided into (8) units. Unit one devotes 10 of its work pages to initial consonant-consonant work. Three worksheets on (s) endings are included. Also work on rhyming, blends, and irregular words. U.(2) devotes about a third to consonant work, a third to sound-symbol matching, and the rest to introductory work on long-short a. U.(4) is devoted almost entirely to work with short vowels. U.(5) and (6) is almost all review with some emphasis on er and ing endings. Unit (7) primarily short vowel work, although new blends are also presented on (3) worksheets. The last unit is review, but does present the y as in piggy and the ee/ea combination.

3. How does the author present his lessons or material.

Every word in every exercise is to be read aloud by the child. Each workbook unit begins with two or three pages which teach the words to be used in the unit. Also included in each workbook unit are two sections called "Orphan" and "Lost Words" and "Let's Have Fun". These are games which utilize the skills previously presented in each unit.

4. How much teacher time is involved?

This material can be used for any number of children although it lends itself particularly well to individual or small group work and may be used in varying time blocks.

5. Size of group for optimal use

6. Time Element

7. Manipulation of materials

8. Accompanying Apparatus

9. Illustrations

Illustrations in the reader are colored, simple drawings with little detail. They appear either at the top or bottom of the page to allow for an even flow of print across the page. Workbook illustrations are simple sketches making use of a small amount of neutral colors.

10. Paper and Printing

The paper is a non-gloss with large print.

11. Manual

There are two separate manuals. One is written for educators and a less complicated edition is provided for parents and laymen.

12. Suggestions

Since this series relies heavily on the skills of structure and configuration it would probably be best used with children who have good visual retention and recall. Children having difficulty in this area would possibly have their reading problem complicated through initial use of this series in remediation. Because of the use of the basic Dolch sight words, books utilizing these words in abundance would be good follow-up materials.

13. Correlated materials

Consonant Pictures for Peg Boards

DIAGRAM #7

MATERIALS ANALYSIS GUIDE CARD: READING

1. Basal
2. Remedial
3. Phonics
4. Individualized
5. Linguistics
6. Readiness
7. Vocabulary
8. Literature
9. Study Skills
10. Hi-Interest Low-Voc.
11. Reading-Language Dev.
12. Speech-Language Dev.
13. Word Analysis
14. Comprehension
15.
16.
17.
18.
19.
20.
21.
22.
23.
24.
25.
26.

DIAGRAM #8

PROBLEM AND FORMAT CARD: READING (PHONICS)

1. Sound Discrimination
2. Initial Consonants
3. Long Vowels
4. Short Vowels
5. Rules
6. Digraphs
7. Diphthongs
8. Endings
9. Blends
10. Regular Double Vowels
11. Irregular Double Vowels
12. Silent Letters
13. General Coverage
14.
15.
16.
17.

A. Workbook

B. Textbook

C. Tape

D. Transparency

E. Filmstrip

F. Slide

G. Duplicated

H. Prepared Cards

I. Games

J. Charts

K. Phonetic Reader

L. Programmed

M. Complete Phonetic Program

N. Inventory

O. Diagnostic Test

P. Record

Q.

DIAGRAM #9

SHELF NO. PHI-PHI.5 R.L. Phonically Graded

FORMAT A,B Check

PROBLEM NO. 18 M.A. 6.5 - 9.5

DESCRIPTIVE ANALYSIS CARD: Reading (Phonics)

TITLE: Easy Road to Reading Improvement (Series)
AUTHOR: Guiliano, William
PUBLISHER: Marand Publishing Co., Inc.

DESCRIPTION: The Easy Road to Reading Improvement consists of two readers with accompanying workbooks. Words containing consonants and long vowels are used predominantly in the first stories. A few non-phonetic words are presented. The author intends his material to be used with elementary school children who wish to improve their reading skills. Words with new phonic elements are introduced every 5 to 15 pages. Each of the 461 basic words is used on at least 8 pages of the reader and on at least 5 pages of the corresponding workbook. Workbooks accompanying the readers contain the same basic words and phonic elements in corresponding units. Child first does a unit in workbook, then corresponding unit in the reader. New words are first taught as sight words. Requires close teacher direction, although some workbook pages can be completed independently of the teacher after proper introduction. Lends itself very well to individual or small group work and can be used in varying amounts of time units. Illustrations in the reader are colored, simple drawings. They appear either at the top or bottom of the page to allow for an even flow of print across the page. Workbook illustrations are simple sketches making use of a small amount of neutral coloring. The print is large on non-gloss paper. There are two separate manuals. One is written for educators and a less complicated edition is provided for parents and laymen.

METHODOLOGY: "Over-learning" Card No. 3
REFERENCES: Albert Harris and Edward Thorndike
EVALUATION: 5
CORRELATED MATERIAL: Consonant Pic. for Peg Boards

DIAGRAM #10

SHELF NO. PHI.1 - PHI.2 R.L. Phonically Graded

FORMAT A, B Check

PROBLEM NO. 2,8,9,.2,.9,20,21 M.A. 6.5 - 9.5

DESCRIPTIVE ANALYSIS CARD: Reading (Phonics)

TITLE: Pat, Lad and the Sleepy Pig
AUTHOR: Guiliano, William
PUBLISHER: Marand Publishing Co., Inc.

DESCRIPTION: Pat, Lad and the Sleepy Pig reader contains 182 basic words. The workbook uses some extra words which are almost exactly like the basic words. The workbook is divided into (8) units. U.(1) devotes 10 of its pages to initial consonant-consonant work. Three worksheets on (s) endings are included. Also work on rhyming, blends, and irregular words. U.(2) is about a third consonant work, third sound - symbol matching, and the rest introductory work on long-short a. U.(4) is devoted almost entirely to work with short vowels. U.(5) and (6) is almost all review with some emphasis on er and ing endings. Unit (7) is primarily short vowel work, although new blends are also presented on 3 worksheets. The last unit is review, but does present the y as in piggy and the ee/ea combination. Every word in every exercise is to be read aloud by the child. Each workbook unit begins by two or three pages which teach the words to be used in the unit. Also included in each workbook unit are two sections called Orphan and Lost Words and Let's Have Fun. These are games which utilize the skills previously presented in each unit. Illustrations in the reader are colored simple drawings. They appear either at the top or across the bottom of the page to allow for an even flow of print across the page. Workbook illustrations are simple sketches making use of a small amount of neutral colors. The print is large on non-gloss paper.

METHODOLOGY: "Over-learning" Card No. 3
REFERENCES: Albert Harris and Edward Thorndike
EVALUATION: 5
CORRELATED MATERIAL: Consonant Pic. for Peg Boards

DIAGRAM #11

SHELF NO. PH.2 - PH.3　　R.L. Phonically Graded

FORMAT A, B　　Check

PROBLEM NO. 2,3,4,6,7,8,9, 14,15,20　　M.A. 6.5 - 9.5

DESCRIPTIVE ANALYSIS CARD: Reading (Phonics)

TITLE: Rednose the Elf
AUTHOR: Guiliano, William
PUBLISHER: Marand Publishing Company

DESCRIPTION: Rednose the Elf presents 279 new words beyond those used in the A reader of this set. Four pages of Unit (1) deal with long and short vowels. Two pages on rhyming. Digraph th is also introduced. In U. (2) the letter J is introduced for the first time. About a third of the unit is dipthong work, the rest digraphs and blends. U. (3) is on blends, endings and more complicated diphthong work, in approximately equal number. In U. (4) the primary emphasis is on blends, although much review material is included. All work in U. (5) is a review of previous work except for the x. Most review is concerned with word families. U. (6) also follows a review pattern but introducesthe letter z and the ook sound. U. (7) presents the qu and deals with ir, er, and ur sounds. The new work in U. (8) deals with the or, au, ew, ou sounds. Workbooks accompanying the readers contain the rest basic words and phonic elements in corresponding units. This material does require close teacher direction. A large percent of the material requires active participation of the teacher. This material lends itself particularly well to individual or small group work. It is easily used in the time usually allowed for a reading program or in even smaller time blocks.

METHODOLOGY: "Over-learning"　　Card No. 3
REFERENCES: Albert Harris and Edward Thorndike
EVALUATION:
CORRELATED MATERIAL: Consonant Pic. for Peg Boards

DIAGRAM #12

E M C INSTRUCTIONAL MATERIALS EVALUATION: TYPES

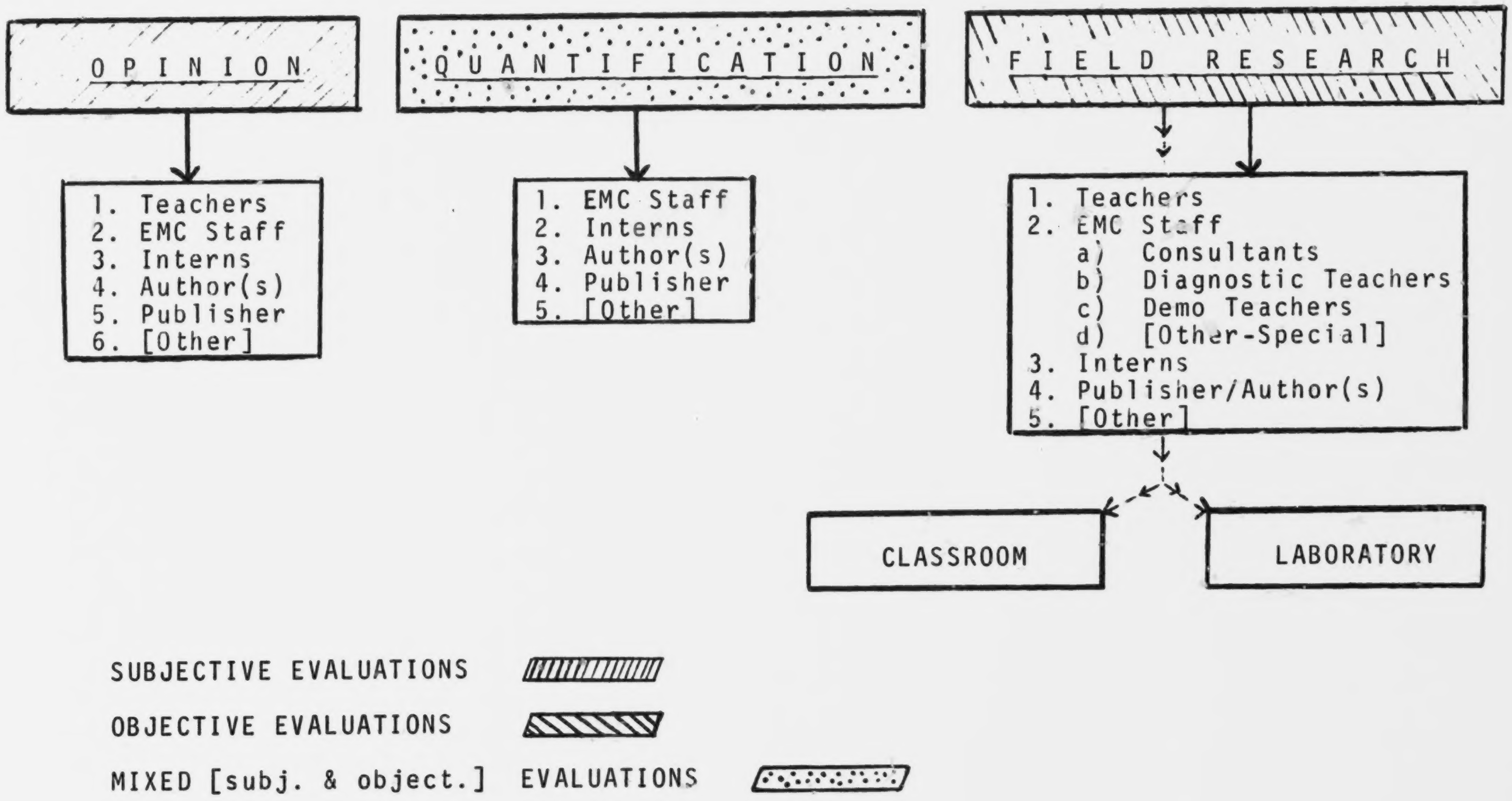


DIAGRAM # 13

MATERIALS EVALUATION FORM

Name____________________ Position____________________ Date____________

Place of employment____

Problem for which material was selected______________________________

MATERIAL____

Briefly state problem for which material was selected________________________

____

(Complete if material was successful)

How effective was material in aiding in remediation?________________________

extremely helpful	very helpful	moderately helpful	little value	no value

For what period of time was material used?__________________________________

1 day	5 days	15 days	other:specify

How much of the material was used? (Be specific) (i.e., pp. 1-12)____________________

How did you modify the original material or method of presentation?_________________

____

____

(Complete if not successful) What factors were important in discontinuing use. (Check those that apply)

_______Student could not use independently. Took too much teacher time.

_______Reading level was too difficult.

_______Student exhibited negative response to format.

_______ Other reasons (Explain)______________________________

DIAGRAM #13 Cont.

PHYSICAL ASPECTS

1. Durability of Material	Highly durable	Adequate	Fairly durable could be improved	Poor durability	Not applicable
2. Packaging of Material	Sturdy packaging	Adequate packaging	Fair packaging could be improved	Poor packaging failing apart	Not applicable
3. General Appearance	Highly attractive	Appealing	Adequate	Unattractive	Not applicable

CONTENT

1. Organization	Highly organized	Well organized	Adequately organized	Fairly organized	Poorly organized
2. Vocabulary Control	Highly controlled	Well controlled	Adequately controlled	Fairly controlled	Poorly controlled
3. Number new concepts	Very high	High	Low	Very Low	Not applicable
4. Supplemental Activities	Many great value	Many offered most appropriate	Adequate offering suitable activities	Few offered most inappropriate	None offered
5. Variety of instructional methods presented	Unlimited variety	Great variety	Adequate variety	Little variety	No variety

DIAGRAM # 13 Cont.

USABILITY

1. Estimated instructional level/s	Pre-sch.	K-1	1-2-3	4-5-6	Jr. High
	Sr. High	Other: Specify			
2. Estimated interest level/s	Pre-sch. K	K-1	1-2-3	4-5-6	Jr. High
	Sr. High	Other: Specify			
3. Effective with	Individuals	Pairs	Small groups	Large groups	Highly flexible in use
4. Teacher's guide	Organized, complete great value	Adequate for successful usage	Somewhat unorganized, Incomplete, of little value	No organization Incomplete No value	None available
5. Teacher involvement	Teacher directed at all times	Teacher introduced; student completes independently inde	Some teacher direction needed in the beginning	Little teacher direction	No teacher direction

Cover inde on bottom of column two

standards by which one makes a judgement.

Quantification may include such characteristics as:

(1) Readability Levels
(2) Sentence Length
(3) Vocabulary Control
(4) Interest Load
(5) Graded Level Placements
(6) Print Size
(7) Color Use
(8) Paper Quality
(9) Layout Quality
(10) Durability

The problem faced in this form of evaluation is the development of standards or formula against which an objective determination can be made.

True, many publishers present this information in their informational literature, but it is generally scant and seldom derived from uniform procedures. An example of the most predominant irregularity is that of readability placements for books described as "High Interest-Low Vocabulary".

The important consideration to be made relative to quantification is the determination of "what to quantify". This can best be decided based on the needs of those using the materials.

Quantification may be done by any number of personnel, but at the EMC it is most often carried out by the staff and by interns fulfilling practicum requirements.

The third type of evaluation in use at the EMC is Field Research. Field Research, as we use the term, has become a more inclusive term than would probably be assigned by others, but we place into this category the evaluation of materials as it is reflected by the change in the child's behavior.

This form of evaluation is generally more time consuming and more costly than other forms. The design of this type of evaluation requires a more stringent control of operational variables, but it contributes more precise results as well as more easily generalized results. This type of evaluation may be done by the teacher in the classroom, the Methods and Materials Consultant/Teacher or others depending upon the purpose of the evaluation.

It is from this form of evaluation of materials that we will gain our most valuable information for prescription writing and programming for children.

As a "side benefit" we see teachers who have become involved in research and evaluation becoming more objective and precise in their teaching as well as being able to make more informed decisions concerning the selection of materials in the future.

Once the evaluation is completed the data is recorded appropriately onto a card system for dissemination purposes. At the same time, the original Descriptive Analysis Card takes on the corresponding evaluation numbers. These numbers refer to an evaluation card which becomes our "prescription bank".

As you can see, we have in effect strengthened our prescriptive program and our knowledge about instructional materials.

EVALUATION PROCEDURES IN MONTGOMERY COUNTY SCHOOLS

by Richard L. Darling, Ph. D.
Director, Department of Instructional Materials
Montgomery County Schools, Rockville, Maryland

A scant six years ago in the Montgomery County Public Schools there were no procedures for evaluation of instructional materials. Perhaps this is an extreme statement, but it is true in the sense that there was no organized and systematic program to identify the best materials to support the county curriculum. No doubt individual schools, within the limitations of their staffs, made a valiant effort. But even then, the number of items of instructional materials appearing on the market made it impossible for a single teacher or even a single school to examine and evaluate all of the potentially useful materials available.

In 1962 a committee called the Superintendent's Advisory Committee on Instructional Materials recommended that a program be initiated. Here are portions of their report.

> Instructional materials are those items which are designed to impart information to the learner in the teaching-learning process. A wide variety of instructional materials is essential for the best instruction. Instructional materials may be consumable or expendable but are generally non-consumable and fairly durable, such as books of various sorts, charts, films, filmstrips, globes, maps, models, magazines, newspapers, pictures, recordings, program materials, slides, specimens, transparencies, workbooks, etc.
>
> Second, the most important objective in all evaluation procedures is to locate and make available for teachers and pupils the most suitable materials that can be found in the various subject areas.
>
> Third, materials should be evaluated by those who are to use them. Group evaluations are generally preferable to individual evaluations. Evaluations are best when they are based upon the actual experience of using the materials in a teaching-learning situation. Instructional materials should be considered in terms of the total curriculum and should be closely coordinated with curriculum revision to assure current and suitable materials.
>
> Fourth, general criteria to be applied when evaluating all types of instructional materials are: appropriateness to a particular curriculum, appropriateness to a particular grade level, authenticity, contribution to learning, quality, good value in terms of purchase price. The above criteria are also to be applied when evaluating free or inexpensive materials. In addition, the following items must also be considered when evaluating instructional materials: characteristics of students relative to their interests, attitudes, experiences, knowledge and skills, the learning pattern which will be most beneficial for the students, the total curriculum of the Montgomery County Public Schools and the specific educational goals of the Montgomery County Public Schools.

As a result of this report, the review and evaluation program was set up within the department of instructional materials, which included other instructional materials operations as well: supervision of the instructional materials programs within the individual schools of the district; a county central instructional materials center, largely a center for inventory of 16 millimeter films and other materials appropriate for central office holdings; a processing center to process library books and other instructional materials; and production facilities for producing new materials.

The first step after the decision had been made to create this service was to establish firm criteria for the evaluation of materials. As a result of a cooperative effort among the departments of supervision and curriculum development, the department of instructional materials, teachers from the classrooms, principals and others, the criteria and procedures were established and published in a document called "Review and Evaluation Procedures for Textbooks and Instructional Materials." This document has gone through several revisions and is distributed in each edition to every teacher, librarian and other educator within the school system. The basis of our program, taking the principles set down by the original advisory committee, is that group evaluation is better than individual evaluation, and as a result, committees were set up to do the bulk of the evaluating of the various instructional materials. There are approximately sixty committees. They are organized by subject and by level.

For example, there is an elementary mathematics evaluation committee, a secondary mathematics evaluation committee; there is a kindergarten through third grade social studies committee, a grade four through six social studies committee. On the secondary level there are committees which devote their efforts to evaluation of materials in specific subjects: a geography committee, an American history committee, and so on.

These committees are made up largely of teachers who volunteer for this service. The Division of Review and Evaluation secures volunteers by public announcement and by asking supervisors, principals and area directors within the system to recommend teachers to serve on these committees. The majority of the members of every committee must be classroom based or school based teachers and the chairman, who is elected by the members, must be a school based teacher. Others who serve on the committees are supervisors in the appropriate subject, librarians, counselors, and some administrators.

The function of the committees is to review and evaluate all of the materials of instruction in the subject at the level at which they work. From their evaluation activities decisions are made as to what materials shall be purchased for the central inventory and/or what materials will be recommended for school purchase. They meet monthly or more frequently, and a schedule is posted on a regular basis in the Division of Review and Evaluation. When the committees come in to evaluate materials they find the materials waiting for them and the necessary equipment there so that they may devote their total effort to the professional task of evaluating materials. Some materials are evaluated on the spot, others, particularly printed materials, are taken away to be evaluated individually, but the committees come back together to discuss the materials and to arrive at final conclusions concerning them.

Textbooks and programmed materials require six signatures to be approved; all others require only three and do not require the signature of the subject supervisor. In addition to the volunteer program which I have just described, (and these committees meet ordinarily after the end of classes on school days) we have a summer program. In Montgomery County Schools 25% of all the professional staff work on a twelve month assignment. During the summer months we have full time committees working on the bulk of the subject areas, devoting either full or part time to the evaluation of materials through the summer. This program, in a sense is not voluntary, because the teachers are being paid to devote their full efforts to the evaluation of materials.

There is one exception which I must mention, to this committee procedure: this relates to trade books. The volume of publication of trade books is so great in any one year that the committees in most subject areas, though not all, are unable to handle the full volume of publications. Therefore, with trade books, we provide for individual evaluation by individuals within the total professional staff of the system, and even provide a combination where one or two professional evaluations from within the system may be coupled with one or two professional evaluations from journals to provide us with the three approving signatures that we require.

This brings me to the Division of Review and Evaluation itself. Having said that the evaluation is done by committees and other individuals in the professional staff of the system, I have indicated that the staff of the Division of Review and Evaluation does not, in itself, do the reviewing. On the contrary, this staff is much too small to hope to encompass the reviewing job to be done. Instead, it is a coordinating staff responsible for the logistics operations in connection with the entire review and evaluation procedure.

A major job of the Division staff is securing materials. They must bring in the materials which committees and individuals will evaluate. They have been averaging for the last four years, thirty thousand items of instructional materials per year. They come in through various ways: automatic samplings from textbook and trade publishers and from some producers of audio-visual materials; the staff may request materials which have been requested in turn by members of the staff of the school system; any member of the professional staff of Montgomery County Public Schools may ask that we bring materials in for evaluation. These requests are given top priority in the entire process. The other way in which the staff gets materials is by meeting with vendors, discussing with them new products, and requesting directly that their new materials be sent in to be evaluated by our staff.

Another function of the Division of Review and Evaluation is coordinating the committee work. I guess I left out the first one: it is organizing the committees. They bring materials before the committees, they schedule the meetings of the committees, and for some types of materials, they make an attempt to bring all the materials of that type together at one time. A good example of that is the map and globe evaluation, which is a one-shot activity in review done just once a year. They bring seventy to one hundred maps and globes together, and that, I believe is something that no teacher and no school could possibly do for itself.

Still another function of the Division of Review and Evaluation is the dissemination of its information. This is done, of course, on an informal basis as teachers and librarians come to the Review and Evaluation Division or phone in or write memorandums requesting information; but it is done in a systematic way through the issuing of lists. The division issues four lists annually: the map and globe list (which reflects the activity performed in that once a year evaluation of maps and globes); the library book lists (which are issued once a year, and with supplements throughout the school year); elementary and secondary textbook lists (which are issued once a year and include thirteen or more titles for every subject at every grade level, so that teachers may select from a rich variety of approved texts); and the fourth list is the list of other instructional materials.

Everything goes onto these four lists from the evaluation process except those items which are for central inventory only. Sixteen millimeter films, for example, are purchased only by the Instructional Materials Center (IMC), another division of the Department of Instructional Materials. The Division of Review and Evaluation simply forwards these evaluations to the Instructional Materials Center where the staff in turn, orders for central inventory.

Still another dissemination activity of the Division of Review and Evaluation is the creation and maintenance of an examination collection, so that teachers, librarians, and supervisors who have received the lists of approved materials may actually examine the materials before they select for their own collections. This examination collection includes a collection of textbooks and library books and a collection of other instructional materials of the non-print type. Every librarian in the school system is provided a minimum of one half day per month to come to the examination center to select materials to take back, to coordinate the evaluation of trade books, or to examine approved materials for selection for their own IMC.

The lists are used in the schools as ordering tools, but they are used in connection with the examination center. I am sure that there are things purchased blindly, but more and more, teachers and librarians, with their lists in hand, come to the Division of Review and Evaluation so that they can examine materials.

Now, this must sound like a fairly rigid program. Let me remind you that we are talking about the 22nd largest school district in the United States and therefore, one where when we approve an item and every school decides to buy it, we are already talking about a lot of money. It sounds rigid, but there are things that keep it from being totally rigid. One of them is a tryout procedure so that new and different instructional materials may be used with children and evaluated with children before they are placed on approved lists. With this procedure, if the school requests permission to try out an item not on the lists, it goes through certain channels; to the Division of Review and Evaluation to make certain that it is not already on an approved list or has not already been rejected for use, and to the appropriate subject supervisor. Then the school is permitted to try it out. They make a written evaluation, which is in turn given to the appropriate committee, which evaluates the materials after the try-out has been completed.

The other exception to our regular procedure is nothing more than an exception. When teachers or principals identify needs with particular children or particular groups of children that cannot be met by materials on approved lists, and they justify the use of some other material, my office has the authority to grant them permission to buy materials not on the approved list.

This is our program, basically, of review and evaluation. I would like to say a little about what its effects have been. One of the things I can say for certain is that by having teachers and supervisors evaluate those things which we purchase for our central inventory, the films and other materials which schools cannot afford to own themselves, we have made sure that we are ordering the materials which teachers want and which teachers will use.

We have had an enormous annual increase in the use of centrally inventoried materials. For the last four years circulation of our 16 millimeter film collections has gone up 25% annually, while the size of our collection has been increasing by 6% to 10%. Our research department has revealed some interesting statistics in connection with this too; I may not have the exact percentage right, but there is someone here who can correct me if I am wrong. In 1961 we knew that with 16 millimeter motion picture films, approximately 9% of classroom time was going to the use of these materials. This fall in one limited group, at least of the teachers, this had increased to a little over 20%.

Another way in which we have been able to find a difference as a result of our review and evaluation program is that the curriculum and instructional materials are more closely related to one another. This is true not only in the classroom and within the school, but in terms of the central office departments of Curriculum and Instructional Materials as well.

The program has developed closer staff relationships. Teachers and librarians work more closely together. Teachers and supervisors, librarians and supervisors work more closely together. The result is a far better staff relationship.

Still another effect is that far more of our staff have a direct say in the selection of materials than would be possible if central office supervisors and other so-called experts were to do all the selecting. With more than 10% of the total professional staff of the school district involved in the evaluation of materials, teachers have a major role in the decision as to what will be used. I believe I can truly say that this program has enhanced the prestige of the teacher in our system. The teachers are the people who decide what materials will be used. They are the ones, basically, who decide which ones ought to be tried out with children before a final decision is made.

With the issuance of lists of materials that have been pre-sifted, with those things which are poor or which are unrelated to our program of studies sifted out, our teachers, librarians and administrators can devote their time and efforts to more important aspects of selection. They can examine materials that have been pre-sorted and can pick those more appropriate to support the program of the school. We can be sure also, that better materials are selected for every school. No longer is it a hit and miss effort but a professional selection from the best materials available for the purpose.

Finally, because these committees examine a great variety of instructional materials, ranging from the text book to the trade book through the whole variety and range of non-print materials in auditory and visual forms, our teachers are better able to find those materials which best suit the need of the children. In other words, materials are brought to bear on the individualization of instruction.

These are a few of the advantages we feel we have gained. We still must improve the program in a number of ways. We must speed it up; I have avoided telling you how long it takes, but we must speed it up because it now takes too long. We must cover a far larger number of materials than we now do, although it is only fair to the staff of the Division of Review and Evaluaton to say that the total of the four lists comes to well over 75,000 individual studies of instructional materials.

We must develop methods which will permit us to involve the children even more than we do in our present method. Now, we use the children when we have materials that are new and different; we think also it would pay if we could try out materials that are more traditional with children, before we make final decisions as to whether or not we should use them. We must devise procedures and secure staff for a more systematic coverage of all materials. Now, we depend upon the samplings provided by producers. We depend upon the requests of teachers and others, and we depend on what the vendors who come in, tell us.

What we need to have is a staff adequate to make a systematic search for materials in every subject of the curriculum and at every level for which we need materials. We must learn how to disseminate more information to our schools. We issue lists now, but the lists are bare-bones lists. They give information on title, author where appropriate, publisher or vendor, price, whether it is black and white, in color, and a few additional things. But we need not only to provide lists, and lists with annotations, but actually to present materials to teachers. We are looking forward in the future, with the development of a television distribution system, to actually showing materials to our teachers, giving them opportunities to hear and see, so that teachers and librarians know every item through this kind of dissemination. There are copyright problems involved, of course, but I think they are insoluble since the intention here is not to substitute for purchase but help schools to learn how to purchase more wisely.

Another dissemination activity to which we must give more attention in the future, though it has certainly not been totally neglected in the past, is in-service education about instructional materials. We must help teachers and librarians learn how to select more effectively, to evaluate the materials they have, to select from those available those they need to have, and to identify those tasks for which no materials exist, so they will know better which materials must be created in the school situation locally.

There are many defects in our program now, but we are identifying them and we are developing plans for their improvement. The millenium, I suspect, will never come, but the program will continue to get better day by day.

Evaluation: Industry's Role

THE CONSUMERS UNION MODEL

by Mr. Morris Kaplan
Technical Director
Consumers Union, Mount Vernon, New York

Although it is not necessary to have a fine film in order to communicate one's ideas, given equal cost and equal energy one would prefer to find the best material on the market. If there were some way that the consumer could find out which item was the best, this would be an improvement over possibly having to choose an inferior item due to lack of information of the available items. It is to this problem that I would like to address myself.

There are characteristics of a good microscope, a good projector or a good tape recorder that are almost independent of the use to which it will be put. No matter how you use it, you want it to be safe. No matter how you use it you would like it to be durable, or convenient to use. And so there are always a host of questions, and they are at a most elementary level. They are not really at all in the area of which we have been talking. If you recognize that, I think we can proceed.

Let me tell about how Consumers Union attempted to answer the question that was raised here earlier--how to go about setting up an agency or process for evaluating products or processes. Who's to do this? Should it be the government in some form or other? Should it be the producers in some form or other? Or how? The solution which Consumers Union arrived at was to do it through a cooperative of the users. If there are 10,000 school districts and if they have any interest at all in products or process evaluation, then perhaps each of them would be prepared to contribute some small sum of money, assign the appropriate investigators to investigate these questions and distribute the information throughout the 10,000 school districts. This in general, is the approach Consumers Union used.

Some thirty-odd years ago a number of people became concerned about the poor quality of products which people were faced with and the choices that they had to make, and how difficult they were. They had the same set of questions I have just mentioned now and to answer them they chose this approach: to establish an independent, non-profit union of consumers. This has certain kinds of implications, and it suggests a single-mindedness of purpose: to provide consumers information on consumer goods and services, independent of a profit motive which could make profit-making more important than this purpose itself.

The fact that we chose a membership or union kind of organization made it a cooperative effort. Members elected a Board of Directors who had no financial interests in the products or any connection with producers or distributors of goods. They were independent. Generally, this Board of Directors consists of scientists and engineers and educators, social workers, and so on. The members of the organization also express their desires through an annual questionnaire on the kinds of things they are interested in. The members pay for it all through their subscription fees and through the purchase of special publications, and although non-members may subscribe, anyone can become a member merely by electing to do so.

Most important was the notion of being independent. That meant that Consumers Union would accept no advertising; it would accept no gifts or donations, or subsidies or grants from any commercial source. It would accept no samples for tests; all of these are bought in the open market. It does not permit the commercial use of Consumers Union test results or Consumers Union's name. It does not permit more than ten copies of any single issue to be purchased by anyone except for educational or non-profit use.

If you are interested in these concepts, then, and find that they might be transferrable to your problem, you will recognize that we start out with subscribers, or members, or users of the service. They elect a Board of Directors who are non-paid and independent. They in turn appoint a director who is the Executive Officer. He hires staff and gets the facilities together. The staff produces a product which is distributed in the form of a published monthly magazine for use of the subscribers.

I will review briefly the general approach that we take to evaluating products. One of the questions we concern ourselves with is, "What products should we evaluate?" For this we set up the Operations Committee which consists of the heads of the departments involved in getting out the reports. There is a marketing group which purchases the samples; there is a library, or information group which feeds in that kind of input; there is the heart of the organization which is the technical department that does the technical product evaluations and then there is the editorial department which concerns itself with distributing this information in a usable form.

The input to the Operations Committee consists of information from an annual questionnaire which we send out to subscribers and from the voluminous correspondence that we get. The magnitude of this correspondence demonstrates the interest that people have in this kind of thing, and I expect that if there were a similar organization in the educational field the response would be similar. There would be lots of people who would be concerned with questions such as: What projector shall I buy? Which will be the easiest to maintain? Which will be the most effective?

We also subscribe to the trade press and read all the trade literature and have lots of similar inputs. The committee decides what projects to undertake and when to publish for timeliness and interest and proper balance of the issues, so that our magazine will be readable.

One of the next problems we concern ourselves with is deciding which brands and models to include. It is rare that we cover the whole field. The Marketing Department determines what products are the most widely available, which have the greatest interest for other reasons. For example, a product may be of interest because of its unusual claims or low cost, or other special features which might be important. This information is put together with a discussion of marketing practices, seasonality of sales, type of outlets through which the products are sold, delivery and installation problems, warranty practices, price information and so on. It assembles manufacturers' specification data. On the basis of its studies it recommends a scope for a project; the scope may be narrow and include a particular narrow category of products or it may be broad and include all kinds of products.

One of the most important questions involves deciding what to test and how to test. I was most impressed with the earlier presentation that discussed this problem and we worry about the same kinds of questions: Are we testing for the right things? Do we know the ambience in which this product is going to be used? And, in the context of that ambience are we evaluating the product properly? But, we have learned over a great many years that even if it is not possible to answer those questions well, there are so many other important questions that one can answer, it is clearly worth the candle to go through this operation and not be hung up by the things we cannot do. We concern ourselves with the things we can do. We can learn a lot by lining up twenty or thirty or forty products, one next to the other. And even if we know nothing more about these products than what we can see, there would be clearly demonstrable differences, sufficiently important to warrant the dissemination of this information. Now we, of course, do have a great deal of technological know-how that has been developed over the years to provide even better answers.

If someone were to press me in terms of the discussion earlier this morning -- Do we know the process well enough? Do we understand all the uses to which these products are going to be put? Does our evaluation answer the question in terms of all of these processes, and wouldn't it be different if we were to answer it for other processes? I really would have great difficulty in answering. But I would argue that what we do is so valuable that one need not be disheartened by the difficulties of the problems that one can raise about such evaluations. If one merely wades in and starts he will find that there is a great deal of extremely useful information that can be obtained and that will help immensely with the decision-making.

What to test for then, and how to test? In terms of consumer products, our library information source supplies the standards and specifications that have been developed for such products all over the world; and if you look at them you realize how elementary, how inadequate, they are. It supplies publications of other studies on the subject, as well as other, more general, references. An engineer, chemist or technologist to whom the project is assigned writes to sources such as the manufacturers themselves, to other test laboratories which concern themselves with this problem, and to others, for information on criteria and methods. He may initiate a questionnaire to ascertain what users consider important. Asking the consumer is a useful kind of exercise, but one ought to know what its serious limitations are.

The sophisticated tester will devise test procedures of his own. There is, of course, the danger of the tyranny of the tester. One of the things he does is test for criteria he knows about. He tests by methods he knows how to use. These are not necessarily the ones that are responsive to the need, but he himself is a captive of the limitations of his craft, and unless he is fully aware of this (and it takes a sophisticated tester to be aware of this) he can mislead you and himself grossly.

Most of the methods in the field have been developed by industry, by the manufacturers, and are industry-oriented. I daresay this is more true in the area of educational materials than it is in the area of consumer goods. The characteristics we try to think about we classify under five headings, and they may be useful to you.

The main heading we refer to is performance, and this has to do with how well the product does the job for which it is designed. If it is a projector, we are interested in how sharp the picture is and characteristics of that sort.

There is a secondary category of characteristics that we refer to as convenience characteristics, and those are: How easy is it to get the product to perform what it is suppossed to do? Are the knobs easily accessible? Are they hard to turn? Is it easy to load and unload?

Then there are a set of considerations under the heading of safety. These have to do with safety for the user, safety for the product that is put into the machine, like the films, safety of the machine itself, and safety of its environment; the table on which it sits, for example.

Another set of characteristics has to do with the durability of the product--that is, how long it is going to last, how trouble free the operation will be, etc.

And finally, we have to deal with economic factors. How costly is it to buy, to operate and to maintain?

These five categories do not include such things as style or appearance, which are also important considerations, but ones we think need to be left to the individual user of the product.

Having gone through this exercise of deciding what to test for and how to test, the project starts. Samples are ordered. We have shoppers, a large number of them around the country, who buy the samples for us at retail and send these samples to our laboratory. This concept is important, it seems to me. Accepting samples from manufacturers is not good practice, for many reasons. For example you may get a sample which is not representative of the product you are interested in evaluating. Also, there is a kind of obligation that you assume when you accept products from manufacturers that you do not want to be under.

Samples are bought and sent to the laboratories. In the laboratories testing equipment is procured, set up, and de-bugged, and then the testing process starts. The samples are subjected to many tests and the data is accumulated. We use a variety of types of testing that may also be interesting to you. We may use laboratory instruments of one sort or another. With that, as with all testing, there is the problem of validation-- how well does a set of test conditions simulate what actually goes on in use? These are not easy questions, more difficult for some products than for others.

We have small panel "use" tests, which again have all kinds of theoretical problems, and are subject to all kinds of criticisms. Yet in practice they turn out to be extremely informative. You can learn a lot about a product by having even a small, unsophisticated group of people use it. The defects which show up even in this simple screening are amazing. Sometimes, one wonders whether the manufacturer even had his product tested by anyone before it went on the market.

Finally, there are elaborate field trials; another technique for evaluating products. These are the most useful, I think, if the field panel is adequately chosen and the test design is appropriate. They are also the most costly, the most complicated, and the most time-consuming, but the easiest to interpret because they involve real life situations.

Then there is the problem of having accumulated a vast amount of data on different characteristics, (sometimes as many as forty or fifty of them). There are twenty or thirty products which one has. And there is a mass of data which needs to be added up some way. It is not enough to merely present all of this information about all of these products, because most users would be confused by all of this. It is not possible to look at twenty times fifty bits of data and try to make some sense out of this. It is rare that one of these twenty products turns out to be good in all respects.

The relevance of data involves weighing of the factors, and it becomes most important to know how the product is going to be used. Some assumptions must be made. We have often been told, and if you are going to attempt to use this approach in solving your problems you will be told by all the producers and by the people who are opposed to this model of product evaluation, that each user is unique and therefore it is not possible to devise a scoring scheme, a way of integrating this data, except in terms of the uniqueness of the individual user. This is a lot of nonsense. Again it is the same kind of sophistication, over-sophistication that is way beyond the level at which we are dealing with these problems. Nobody wants a product that's unsafe. Nobody wants a product that will fall apart after a week of use. Everyone wants a product that performs well even at the most elementary level. The light must go on when you press the button. Now if you limit yourself to that most elementary kind of evaluation, you will still find clear and important distinctions. You will be able to reduce these twenty brands to three or four or five that have the virtues you are looking for; the others will all fall by the wayside for one or another of the kinds of deficiencies I have talked about that are so serious that no matter how the product is going to be used it will be deficient. So it is possible, I think, on a crude level, but still a very useful one, to devise rating schemes, scoring schemes of various types and a number of similar approaches.

Next we go through an elaborate process involving checking of the data. It is essential if you are to retain the credibility of the whole operation that you check the data. Credibility is a very important word in this union of consumers. If there is any question about your doing your work well, if there is any question about your lack of bias, your independence, then the whole operation falls by the wayside. What has made Consumers Union as unique and as successful as it is, is that it has been possible over the years to maintain this independence and reliability.

Finally, there is the writing of a technical report which is re-written by our editorial department; nothing gets published unless the engineers who have done the work are satisfied with the accuracy, tone, and emphasis of the published report. This is very important. This is not what happens in industry, even where industry does do its own product evaluations. The advertising or sales people are the ones usually who have the final word about what goes out to the user.

This, in very rough outline, reviews the kinds of things that we do. Our evaluations are comparative rather than absolute. This makes life a good deal easier, and for most purposes, that is the real life situation-- the choice is among the existing products and you might end up saying that none of them is any good and that therefore one ought not to buy any of them. But this is beyond the kinds of questions I am raising.

Actually, it is rare that all products are unsafe, although we have had situations of that sort, and you might decide, if you know that, not to buy any of them and find some other way to achieve the objective. But the usual situation is that you are going to buy a projector anyway; you just have to have a projector, and the problem is which of the twenty or thirty or forty you are going to buy. Therefore, comparative evaluations are the most meaningful and useful ones.

INDUSTRY'S ROLE IN EVALUATION AND EDUCATION

by David R. Dorsett
Regional Consultant, Creative Playthings. Odenton, Maryland

Historically there has been a serious gap between the classroom teacher and the educational industry. This gap has caused undue hardships for both the educator and the producers of materials and it is on this problem that my thinking has focused in the evaluation of instructional materials.

It is my belief that the responsible service to the customer indicates a more active participation on the part of industry in all phases of education. Just the fact that industry is represented here today and is being called upon to participate in more and more high level conferences dealing with educational concerns is an indication of industry's active role in helping to set educational standards.

Specifically, industry is becoming more and more engaged in active dialogue with its customers. My company, as well as many others, has felt the need to communicate directly with teachers. We developed a program of desk-top manipulatives for the elementary grades, and educators manifested a keen interest in this program from its inception. Conferences with leading educators were the guiding factors influencing the research and ultimate development of this program. Yet a very real need was felt to communicate directly with those who were to put the program into use . . . those whose daily task was working with children.

A series of workshops were held, and are being held, to promote this dialogue between educators and industry. Response to the program has been overwhelming. In attempting to understand the teacher's problems, needs, and goals, we can go to greater lengths to be of use to schools and children.

Because of industry's expensive experience with the development of educational programs, it is a fact that insights gained in content and pedagogy are useful to the educator. There appears to be a greater need and a greater willingness to have industry comment on issues, evaluate, innovate and communicate its findings to the professions.

Industry is beginning to share in the responsibility of materials selection. It has a responsibility to be sure that what is being presented for consumption has merit. Our new catalogue will contain several sections which deal exclusively with curriculum and materials innovations based upon contemporary research.

Needless to say, it is necessary to be sure that both industry and education are speaking the same language. The greater the involvement in and understanding of each, in all phases of education, the less chance of failure. Each should be made aware of the other's problems and consequently move toward the same basic goals. Many of the suggested guidelines for the evaluation of instructional materials specifically involve a role for industry. For example:

1. What is the specific purpose of this material?

Every product presented to the educator should be accompanied by concise statements concerning possible uses of the particular device. This does not mean that we need necessarily limit the use or prescribe treatments in the exact dosages but we do have the responsibility to inform the educator not only that this product exists, but why it exists.

2. Does this material fit the needs of the particular group who will be using it? If not, how can it be adapted to suit this need?

The educator is closest to the needs of his particular group of children, and the materials produced must be suited to his requirements. Industry as well as the teacher has a real interest in creating the correct materials, in constantly improving upon materials now in use, and in adding to existing programs. Involving more educators directly in the research and development sections of industry is a prerequisite for the growth of better instructional programs and better manipulatives to motivate, instruct and guide children.

3. Is what is being presented accurate and sound?

A specific evaluation of this type is often made in the classroom itself. By observation of materials in use in a real situation the theory, the philosophy, the "feelings" about the usefulness of a certain piece of material can be more readily determined. Industry needs to take a more active part in the classroom itself through observation, through field-testing and even through direct involvement in the process of teaching.

4. What are the physical qualities of the material?

As well as a need for standards in the curricular value of materials there is also a need to construct materials to a certain quality specification. Their attractiveness, their sturdiness, the texture of the object itself may either motivate a child or frustrate him.

Poor design needs to be called to the attention of industry. Most industry welcomes suggestions for variations in design so as to improve its product. This type of communication and action serves not only the interest of the industry but the interest of the consumer.

5. Does this material offer variety to the program?

Instructional materials which can be used in great varieties of ways are of greatest use to the teacher.

6. Are the teacher's manuals which accompany the materials sufficient so as to suggest additional uses, pose new problems, and suggest correlations with other sources?

A set of well defined and clearly stated objectives should accompany a teacher's manual. The purpose of the manual need not be one of limitation or specification. Its function is one of an introduction to the program. The manual should give practical guidance by offering a variety of plans and yet foster teacher creativity by allowing for deviation from the suggested procedures. Information related in a manual should be factual; so factual that the educator is aware of its intended usage and not simply the opinion of the author or the producer. Good manuals suggest correlations with other materials, other texts, library references, films and filmstrips and other media. It is certainly a responsibility of industry to be sure that all printed literature concerning a product meets the standards set by the product itself and that this literature is a statement of fact and to the best of their knowledge is reliable.

7. To what extent will these materials accomplish the specific goals of the educator?

Many classrooms have been overwhelmed with volumes of new materials, texts, audio-visual aids, three-dimensional manipulatives and electronic equipment which were thought to be a cure-all for all of their problems. We seem to be in a "cooling-off" stage at this time--waiting for research to support or deny that which was innovative in the past decade.

8. Is the evaluation process itself limited to certain periods of time?

In summary it is apparent that any evaluation must be done in relation to the product of evaluation--the individual child. Another recent NEA publication states: "Provisions of buildings, equipment, administration, materials, personnel, and services are useful only as they help children develop. The preparation of the teacher, his selection of materials, and everything he does to meet the demands of the teaching day are good or not good in terms of the results in children. Success is measured not by how many days are good, but by how many children are reached."